```
      DO 120 JJ=1,II
      K = II - 39
      IF ((JJ.GE.4).AND.(JJ.LE.37)) THEN
      I = JJ - 3
      DO 110 J=1,4
      DO 110 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0   GO TO 110
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ)   A(II,JJ)  + WW*(1 WW)
110    CONTINUE
      A(JJ,II)   A(II,JJ)
      ELSE IF ((JJ.GE.38).AND.(JJ.LE.40)) THEN
      J = JJ   36
      DO 115 I=1,34
      DO 115 L 1,31
      IF (X1(I,J,K,L) .EQ. 0.0) GO TO 115
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ) = A(II,JJ) + WW*(1 WW)
115   CONTINUE
      A(JJ,II) = A(II,JJ)
      ELSE
      DO 117 I=1,34
           DO 117 J=1,4
           DO 117 L=1,31
           IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 117
           WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
           IF ( JJ .EQ. 2 ) THEN
           A(II,JJ) = A(II,JJ) + X1(I,J,K,L)*WW*(1-WW)
           ELSE IF ( JJ. EQ. 3 ) THEN
           A(II,JJ) = A(II,JJ) + X2(I,J,K,L)*WW*(1-WW)
           ELSE IF ((JJ.EQ.1).OR.(JJ.EQ.II)) THEN
           A(II,JJ) = A(II,JJ) + WW*(1-WW)
           ELSE
           A(II,JJ) = 0.
           ENDIF
117         CONTINUE
      A(JJ,II) = A(II,JJ)
      ENDIF
120   CONTINUE
      DO 140 II=44,77
      DO 140 JJ=1,II
      I = II - 43
      IF ((JJ.GE.4).AND.(JJ.LE.37)) THEN
      DO 130 J=1,4
      DO 130 K=1,4
      DO 130 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 130
      IF (JJ .NE. (II-40)) THEN
      A(II,JJ) = 0.
      ELSE
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ) = A(II,JJ) + WW*(1-WW)*X3(I,J,K,L)
      ENDIF
130   CONTINUE
```

```
      A(JJ,II) = A(II,JJ)
      ELSE IF ((JJ.GE.38).AND.(JJ.LE.40)) THEN
      J = JJ - 36
      DO 135 K=1,4
      DO 135 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0) GO TO 135
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ) = A(II,JJ) + WW*(1-WW)*X3(I,J,K,L)
135   CONTINUE
      A(JJ,II) = A(II,JJ)
      ELSE IF ((JJ.GE.41).AND.(JJ.LE.43)) THEN
      K = JJ - 39
      DO 137 J=1,4
      DO 137 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0) GO TO 137
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ) = A(II,JJ) + WW*(1-WW)*X3(I,J,K,L)
137   CONTINUE
      A(JJ,II) = A(II,JJ)
      ELSE
      DO 139 J=1,4
            DO 139 K=1,4
            DO 139 L=1,31
            IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 139
            WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
            IF ( JJ .EQ. 2 ) THEN
            A(II,JJ)              =              A(II,JJ)
     X1(I,J,K,L)*WW*(1-WW)*X3(I,J,K,L)
            ELSE IF ( JJ. EQ. 3 ) THEN
            A(II,JJ)              =              A(II,JJ)
     X2(I,J,K,L)*WW*(1 WW)*X3(I,J,K,L)
            ELSE IF (JJ.EQ.1) THEN
            A(II,JJ) = A(II,JJ) + WW*(1 WW)*X3(I,J,K,L)
       ELSE IF (JJ.EQ.II) THEN
       A(II,JJ) = A(II,JJ) + (X3(I,J,K,L)**2)*(WW*(1-WW))
            ELSE
            A(II,JJ) = 0.
            ENDIF
139          CONTINUE
      IF ( JJ. EQ. II ) A(II,JJ) = A(II,JJ) + 1/SS2
      A(JJ,II) = A(II,JJ)
      ENDIF
140   CONTINUE
      RETURN
      END
      FUNCTION TWOW (X,X1,X2,X3,I,J,K,L,N)
      DIMENSION X1(34,4,4,31),X2(34,4,4,31),X3(34,4,4,31)
      DIMENSION X(N)
      IF (J. EQ. 1) GO TO 95
      IF (K.EQ.1) GO TO 100
85    W = X(1)+X(2)*X1(I,J,K,L)+X(3)*X2(I,J,K,L)+X(I+3)+X(J+36)+
     1X(K+39)+X(I+43)*X3(I,J,K,L)
      TWOW= EXP(W)/(1+EXP(W))
      GO TO 110
```

```
95     IF (K.EQ.1) GO TO 105
       W = X(1) + X(2)*X1(I,J,K,L) + X(3)*X2(I,J,K,L) + X(I+3) +
      1X(K+39) + X(I+43)*X3(I,J,K,L)
       TWOW= EXP(W)/(1+EXP(W))
       GO TO 110
100    W=X(1)+X(2)*X1(I,J,K,L)+X(3)*X3(I,J,K,L)+X(I+3)+X(J+36)+
      1X(I+43)*X3(I,J,K,L)
       TWOW = EXP(W)/(1+EXP(W))
       GO TO 110
105    W=X(1)+X(2)*X1(I,J,K,L)+X(3)*X2(I,J,K,L)+X(I+3)+X(J+43)*
      1X3(I,J,K,L)
       TWOW=EXP(W)/(1+EXP(W))
110    RETURN
       END
```

```
****************
*   MODEL  1   *
****************
```

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
NL, LO.	11	7.92	0.7204	6.40	0.8456

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LRPP	NEW 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	0.0	32.0	0.0	0.0	32.0
	.400000	5.7	7.9	0.0	1.4	15.0
	.600000	1.7	8.0	1.1	1.1	12.0
	.800000	3.2	9.0	0.5	0.3	13.0
	LAST	2.4	2.2	0.7	1.7	7.0
	TOTAL	13.0	59.1	2.4	4.5	79.0
YES	.200000	0.0	1.0	0.0	0.0	1.0
	.400000	2.3	3.1	0.0	0.6	6.0
	.600000	1.3	6.0	0.9	0.9	9.0
	.800000	8.8	25.0	1.5	0.7	36.0
	LAST	51.6	47.8	16.3	36.3	152.0
	TOTAL	64.0	82.9	18.6	38.5	204.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.0	0.0	0.0	0.0
	.400000	0.1	0.1	0.0	-0.4
	.600000	0.2	-0.4	-0.1	0.8
	.800000	1.0	-0.3	-0.7	-0.5
	LAST	-0.9	0.5	0.3	0.3
YES	.200000	0.0	0.0	0.0	0.0
	.400000	-0.2	-0.1	0.0	0.6
	.600000	-0.3	0.4	0.2	-0.9
	.800000	-0.6	0.2	0.4	0.3
	LAST	0.2	-0.1	-0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.0	0.0	0.0	0.0
	.400000	0.2	0.1	0.0	-0.2
	.600000	0.3	-0.3	0.1	0.8
	.800000	1.0	-0.3	-0.8	-0.4
	LAST	-0.8	0.6	0.4	0.4
YES	.200000	0.0	0.2	0.0	0.0
	.400000	-0.0	0.0	0.0	0.6
	.600000	-0.1	0.5	0.3	-1.1
	.800000	-0.5	0.2	0.5	0.4
	LAST	0.2	-0.1	-0.0	-0.0

BL.LO.

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LRPP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	0.0	7.8	0.0	0.0	0.0	21.3	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	3.6	0.0	0.0	0.0	7.1	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	2.3	0.0	0.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0
	.800000	1.9	1.1	0.3	0.3	0.5	0.3	3.4	0.0	0.0	0.3	0.0	0.0
	LAST	0.8	0.3	0.0	0.0	0.4	0.1	0.2	0.0	0.0	0.2	0.1	0.3
	TOTAL -	2.6	1.3	13.9	0.3	0.9	0.4	38.4	0.0	0.0	0.5	0.1	0.3
YES	.200000	0.0	0.0	0.2	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	1.4	0.0	0.0	0.0	2.9	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	1.7	0.0	0.0	0.0	4.7	0.0	0.0	0.0	0.0	0.0
	.800000	5.1	2.9	0.7	0.7	1.5	0.7	9.6	0.0	0.0	0.7	0.0	0.0
	LAST	17.2	5.7	0.0	1.0	7.6	2.9	3.8	1.0	1.0	4.8	1.9	5.7
	TOTAL -	22.4	8.7	4.1	1.7	9.1	3.6	21.6	1.0	1.0	5.5	1.9	5.7

OUT	LRPP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	2.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	400000	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.3	0.3	0.0	0.0	1.1	0.3	0.3	0.0	0.0	0.3	0.5
	LAST	0.0	0.4	0.0	0.5	0.1	0.5	0.0	0.2	0.4	0.0	0.1	0.0
	TOTAL -	8.4	0.7	0.3	0.5	0.1	1.6	0.3	0.4	0.4	0.0	0.4	0.5
YES	.200000	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	[illegible]
	400000	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	[illegible]
	600000	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	[illegible]
	800000	0.0	0.7	0.7	0.0	0.0	2.9	0.7	0.7	0.0	0.0	0.7	[illegible]
	LAST	1.0	9.6	0.0	10.5	2.9	11.5	0.0	3.8	9.6	1.0	2.9	[illegible]
	TOTAL -	3.6	10.3	0.7	10.5	2.9	14.4	0.7	4.6	9.6	1.0	3.6	1.5

OUT	LRPP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.5	0.0	0.3	0.3	0.0	0.5	0.0	0.0	0.0	0.3
	LAST	0.2	0.3	0.1	0.0	0.0	0.6	0.2	0.1	0.3	0.0	0.1	0.3
	TOTAL -	0.2	0.3	0.6	0.6	0.3	0.8	0.2	0.6	0.3	0.0	0.1	0.5
YES	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	1.5	0.0	0.7	0.7	0.0	1.5	0.0	0.0	0.0	0.7
	LAST	3.8	5.7	1.9	0.0	0.0	12.4	4.8	1.9	5.7	1.0	1.9	5.7
	TOTAL -	3.8	5.7	3.4	0.4	0.7	13.2	4.8	3.4	5.7	1.0	1.9	6.5

OUT	LRPP	BRN 47.0000	48.0000	TOTAL
NO	200000	0.0	0.0	32.0
	400000	0.0	0.0	15.0
	600000	1.7	0.0	12.0
	800000	0.3	0.0	13.0
	LAST	0.0	0.1	7.0
	TOTAL	2.0	0.1	79.0
YES	200000	0.0	0.0	1.0
	400000	0.0	0.0	6.0
	600000	1.3	0.0	9.0
	800000	0.7	0.0	36.0
	LAST	0.0	1.9	152.0
	TOTAL	2.0	1.9	204.0

OUT	LRPP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10 0000	11 0000	12.0000	14.0000	15.0000	17.0000	18 0000
NO	.200000	0.0	0.0	-0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.2	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.3	0.0	0.0	0.0	-0.9	0.0	0.0	0.0	0.0	0.0
	.800000	0.1	0.9	-0.5	-0.5	-0.7	-0.5	0.3	0.0	0.0	-0.5	0.0	0.0
	LAST	0.2	-0.5	0.0	-0.2	1.1	-0.4	-0.4	-0.2	-0.2	-0.5	-0.3	-0.5
YES	.200000	0.0	0.0	1.5	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	-0.4	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	-0.5	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0
	.800000	-0.1	-0.5	0.3	0.3	0.4	0.3	-0.2	0.0	0.0	0.3	0.0	0.0
	LAST	-0.1	0.1	0.0	0.0	-0.2	0.1	0.1	0.0	0.0	0.1	0.1	0.1

OUT	LRPP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	-0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	-0.5	-0.5	0.0	0.0	-0.1	-0.5	1.4	0.0	0.0	-0.5	0.6
	LAST	-0.2	0.8	0.0	-0.7	-0.4	0.6	0.0	-0.4	-0.7	-0.2	2.4	0.0
YES	.200000	-0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	-0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.2	0.3	0.0	0.0	0.0	0.2	-0.9	0.0	0.0	0.2	-0.4
	LAST	0.0	-0.2	0.0	0.1	0.1	-0.1	0.0	0.1	0.1	0.0	-0.5	0.0

OUT	LRPP	BRN 35 0000	36 0000	37 0000	38 0000	39 0000	40 0000	41 0000	42 0000	43 0000	44 0000	45 0000	46 0000
NO	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.6	0.0	-0.5	-0.5	0.0	0.6	0.0	0.0	0.0	-0.5
	LAST	-0.4	-0.5	-0.3	0.0	0.0	0.6	-0.5	-0.3	-0.5	-0.2	-0.3	[illegible]
YES	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	-0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	-0.4	0.0	0.3	0.3	0.0	-0.4	0.0	0.0	0.0	0.3
	LAST	0.1	0.1	0.1	0.0	0.0	-0.1	0.1	0.1	0.1	0.0	0.1	-0.3

OUT	LRPP	BRN 47.0000	48.0000
NO	.200000	0.0	0.0
	.400000	0.0	0.0
	.600000	0.2	0.0
	.800000	-0.5	0.0
	LAST	0.0	-0.3
YES	.200000	0.0	0.0
	.400000	0.0	0.0
	.600000	-0.3	0.0
	.800000	0.3	0.0
	LAST	0.0	0.1

****
* MODEL 1 *
****

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
LO, AL	38	55.81	0.0311	71.44	0.0008

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LRPP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	12.6	1.9	8.7	2.9	2.9	1.0	0.0	1.0	0.0	0.0	1.0	0.0
	.400000	2.1	4.3	2.9	2.1	0.0	1.4	0.0	0.7	0.0	0.0	0.7	0.7
	.600000	0.6	5.1	2.9	1.7	0.0	0.6	0.0	0.0	0.0	0.6	0.0	0.6
	.800000	0.3	3.2	4.2	2.7	1.1	0.5	0.0	0.3	0.3	0.0	0.0	0.5
	LAST	0.3	1.9	2.2	0.9	0.7	0.1	0.2	0.2	0.0	0.1	0.0	0.3
	TOTAL -	15.8	16.5	20.9	10.3	4.7	3.6	0.2	2.1	0.3	0.7	1.7	2.1
YES	.200000	0.4	0.1	0.3	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.9	1.7	1.1	0.9	0.0	0.6	0.0	0.3	0.0	0.0	0.3	0.3
	.600000	0.4	3.9	2.1	1.3	0.0	0.4	0.0	0.0	0.0	0.4	0.0	0.4
	.800000	0.7	8.8	11.8	7.3	2.9	1.5	0.0	0.7	0.7	0.0	0.0	1.5
	LAST	5.7	42.1	47.8	20.1	16.3	2.9	3.8	3.8	1.0	2.9	0.0	5.7
	TOTAL -	8.2	56.5	63.1	29.7	19.3	5.4	3.8	4.9	1.7	3.3	0.3	7.9

OUT	LRPP	AMTTN TOTAL
NO	.200000	32.0
	.400000	15.0
	.600000	12.0
	.800000	13.0
	LAST	7.0
	TOTAL	79.0
YES	.200000	1.0
	.400000	6.0
	.600000	9.0
	.800000	36.0
	LAST	152.0
	TOTAL	204.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.1	0.0	-0.2	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.6	-0.1	-0.5	0.6	0.0	-1.2	0.0	0.3	0.0	0.0	0.3	0.3
	.600000	0.6	0.4	-1.1	0.2	0.0	-0.8	0.0	0.0	0.0	0.6	0.0	0.6
	.800000	-0.5	-0.1	-0.1	-1.0	1.9	-0.7	0.0	-0.5	1.4	0.0	0.0	0.6
	LAST	1.4	-1.4	-0.8	1.1	-0.9	5.1	-0.4	-0.4	-0.2	2.4	0.0	-0.5
YES	.200000	-0.6	-0.2	1.4	-0.3	-0.3	-0.2	0.0	-0.2	0.0	0.0	-0.2	0.0
	.400000	-0.9	0.2	0.8	-0.9	0.0	1.9	0.0	-0.5	0.0	0.0	-0.5	-0.5
	.600000	-0.7	-0.4	1.3	-0.3	0.0	0.9	0.0	0.0	0.0	-0.7	0.0	-0.7
	.800000	0.3	0.1	0.1	0.6	-1.1	0.4	0.0	0.3	-0.9	0.0	0.0	-0.4
	LAST	-0.3	0.3	0.2	-0.2	0.2	-1.1	0.1	0.1	0.0	-0.5	0.0	0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.2	0.2	-0.2	0.2	0.2	0.2	0.0	0.2	0.0	0.0	0.2	0.0
	.400000	0.6	-0.0	-0.4	0.6	0.0	-1.6	0.0	0.5	0.0	0.0	0.5	0.5
	.600000	0.6	0.5	-1.1	0.3	0.0	-0.8	0.0	0.0	0.0	0.6	0.0	0.6
	.800000	-0.4	0.0	-0.0	-1.0	1.4	-0.8	0.0	-0.4	1.0	0.0	0.0	0.6
	LAST	1.0	-2.0	-0.7	1.0	-1.0	1.9	-0.3	-0.3	-0.1	1.2	0.0	-0.4
YES	.200000	-0.6	-0.1	1.0	-0.2	-0.2	-0.1	0.0	-0.1	0.0	0.0	-0.1	0.0
	.400000	-1.1	0.3	0.8	-1.1	0.0	1.3	0.0	-0.5	0.0	0.0	-0.5	-0.5
	.600000	-0.6	-0.3	1.1	-0.1	0.0	0.8	0.0	0.0	0.0	-0.6	0.0	-0.6
	.800000	0.4	0.1	0.1	0.6	-1.2	0.5	0.0	0.4	-1.0	0.0	0.0	-0.2
	LAST	-0.2	0.3	0.2	-0.2	0.2	-1.1	0.2	0.2	0.2	-0.4	0.0	0.2

```
****************
*   MODEL  1   *
****************
```

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
LO, AL.	12	12.09	0.4387	13.34	0.3448

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LRPP	AMT 250.000	500.000	750.000	LAST	TOTAL
NO	.200000	25.2	2.9	1.9	1.9	32.0
	.400000	9.3	5.0	0.0	0.7	15.0
	.600000	9.7	2.3	0.0	0.0	12.0
	.800000	7.4	3.2	1.9	0.5	13.0
	LAST	5.4	0.8	0.4	0.4	7.0
	TOTAL	57.1	14.2	4.1	3.6	79.0
YES	.200000	0.8	0.1	0.1	0.1	1.0
	.400000	3.7	2.0	0.0	0.3	6.0
	.600000	7.3	1.7	0.0	0.0	9.0
	.800000	20.6	8.8	5.1	1.5	36.0
	LAST	117.6	18.2	7.6	8.6	152.0
	TOTAL	149.9	30.8	12.9	10.4	204.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.0	0.1	0.0	0.0
	.400000	0.2	-0.4	0.0	0.3
	.600000	-0.2	0.5	0.0	0.0
	.800000	0.2	-0.1	0.1	-0.7
	LAST	-1.5	2.4	1.1	1.0
YES	.200000	0.2	-0.3	-0.2	-0.2
	.400000	-0.4	0.7	0.0	-0.5
	.600000	0.3	-0.5	0.0	0.0
	.800000	-0.1	0.1	-0.1	0.4
	LAST	0.3	-0.5	-0.2	-0.2

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	AMT 250.000	500.000	750.000	LAST
NO	.200000	0.0	0.2	0.2	0.2
	.400000	0.3	-0.3	0.0	0.5
	.600000	-0.2	0.5	0.0	0.0
	.800000	0.3	0.0	0.2	-0.8
	LAST	-1.6	1.6	0.9	0.8
YES	.200000	0.4	-0.2	-0.1	-0.1
	.400000	-0.3	0.7	0.0	-0.5
	.600000	0.3	-0.4	0.0	0.0
	.800000	-0.1	0.1	0.0	0.5
	LAST	0.3	-0.5	-0.1	-0.1

Table A-6.5

Classification Results

		Classified		
Observed	% correct	No	Yes	Total
Logit Regression (Fixed Effects) (LRFE) - 282 cases.				
No	89.74	70	8	78
Yes	91.67	17	187	204
Total	91.13	87	195	282
correct		80.46	95.90	

Table A-7
Results for Best Models.

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	75.86	66	21	87
	Yes	91.27	22	230	252
	Total	87.32	88	251	339
	% correct		75.00	91.63	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	89.87	71	8	79
	Yes	92.65	15	189	204
	Total	91.87	86	197	283
	% correct		82.56	95.94	
C.	Logit Regression (Mixed Effects) (LRME)				
	No	72.15	57	22	79
	Yes	94.12	12	192	204
	Total	87.99	69	214	283
	% correct		82.61	89.72	

Table A-7.1
Best Model LDA on sample of 283 cases

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	70.89	56	23	79
	Yes	94.61	11	193	204
	Total	87.99	67	216	283
	% correct		83.58	89.35	
B.	LDA* (adjusted for cases and effects)				
	No	78.48	62	17	79
	Yes	91.18	18	186	204
	Total	87.63	80	203	283
	% correct		77.50	91.63	

Table A-7.2
Comparison of Naive Models

	Observed	% correct	Classified No	Classified Yes	Classified Total
Naive Model					
	No	0.00	0	87	87
	Yes	100.00	0	252	252
	Total	74.34	0	339	339
	% correct			74.34	
Challenging Naive Model					
	No	43.67	38	49	87
	Yes	89.29	27	225	252
	Total	77.58	65	274	339
	% correct		58.46	82.12	
Special Branch Naive Model					
	No	77.01	67	20	87
	Yes	84.52	39	213	252
	Total	82.60	106	233	339
	% correct		63.21	91.42	
Naive Model Based on Branches					
	No	78.16	68	19	87
	Yes	84.52	39	213	252
	Total	82.89	107	232	339
	% correct		63.55	91.81	

TABLE A-8

Table A-8

Holdout Analysis Number 1

Observed	% correct	No	Classified Yes	Total
A. Linear Discriminant Analysis (LDA)				
No	82.98	39	8	47
Yes	90.76	11	108	119
Total	88.55	50	116	166
B. Logit Regression (Fixed Effects) (LRFE)				
No	100.00	43	0	43
Yes	97.89	2	93	95
Total	98.55	45	93	138
C. Logit Regression (Mixed Effects) (LRME)				
No	74.42	32	11	43
Yes	98.95	1	94	95
Total	91.30	33	105	138
D. LDA* (adjusted for cases and effects)				
No	90.32	28	3	31
Yes	96.77	3	90	93
Total	95.16	31	93	124

Table A-8.1

Holdout Analysis Number 1

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	42.50	17	23	40
	Yes	93.98	8	125	133
	Total	82.08	25	148	173
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	47.22	17	19	36
	Yes	94.50	6	103	109
	Total	82.76	23	122	145
C.	Logit Regression (Mixed Effects) (LRME)				
	No	41.67	15	21	36
	Yes	98.17	2	107	109
	Total	84.14	17	128	145

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	45.83	22	26	48
	Yes	90.09	11	100	111
	Total	76.73	33	126	159

Table A-8

Holdout Analysis Number 2

Observed		% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	77.50	31	9	40
	Yes	89.19	12	99	111
	Total	86.09	43	108	151
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	94.44	34	2	36
	Yes	94.38	5	84	89
	Total	94.40	39	86	125
C.	Logit Regression (Mixed Effects) (LRME)				
	No	80.56	29	7	36
	Yes	91.01	8	81	89
	Total	88.00	37	88	125

D.	LDA* (adjusted for cases and effects)				
	No	97.30	36	1	37
	Yes	92.38	8	97	105
	Total	93.66	44	98	142

Table A-8.1

Holdout Analysis Number 2

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	48.94	23	24	47
	Yes	89.36	15	126	141
	Total	79.26	38	150	188
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	62.79	27	16	43
	Yes	74.78	29	86	115
	Total	71.52	56	102	158
C.	Logit Regression (Mixed Effects) (LRME)				
	No	69.77	30	13	43
	Yes	93.04	8	107	115
	Total	86.71	38	120	158

	Observed	% correct	No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	71.43	30	12	42
	Yes	80.81	19	80	99
	Total	78.01	49	92	141

Table A-8

Holdout Analysis Number 3

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	78.72	37	10	47
	Yes	92.20	11	130	141
	Total	88.83	48	140	188
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	90.70	39	4	43
	Yes	95.65	5	110	115
	Total	94.30	44	114	158
C.	Logit Regression (Mixed Effects) (LRME)				
	No	67.44	29	14	43
	Yes	95.65	5	110	115
	Total	87.97	34	124	158

D.	LDA* (adjusted for cases and effects)				
	No	84.62	33	6	39
	Yes	95.92	4	94	98
	Total	92.70	37	100	137

Table A-8.1

Holdout Analysis Number 3

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	70.00	28	12	40
Yes	85.59	16	95	111
Total	81.46	44	107	151
B. Logit Regression (Fixed Effects) (LRFE)				
No	63.89	23	13	36
Yes	80.90	17	72	89
Total	76.00	40	85	125
C. Logit Regression (Mixed Effects) (LRME)				
No	75.00	27	9	36
Yes	87.64	11	78	89
Total	84.00	38	87	125

Observed	% correct	No	Yes	Total
D. LDA* (adjusted for cases and effects)				
No	62.5	25	15	40
Yes	82.08	19	87	106
Total	76.71	44	102	146

Table A-8

Holdout Analysis Number 4

	Observed	% correct	No	Classified Yes	Total
A.	Linear Discriminant Analysis (LDA)				
	No	85.00	34	6	40
	Yes	85.82	19	115	134
	Total	85.63	53	121	174
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	85.71	30	5	35
	Yes	95.58	5	108	113
	Total	93.24	35	113	148
C.	Logit Regression (Mixed Effects) (LRME)				
	No	60.00	21	14	35
	Yes	92.03	9	104	113
	Total	84.46	30	118	148

D.	LDA* (adjusted for cases and effects)				
	No	86.67	39	6	45
	Yes	94.00	7	109	116
	Total	91.93	46	115	161

Table A-8.1

Holdout Analysis Number 4

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	57.45	27	20	47
	Yes	93.22	8	110	118
	Total	83.03	35	130	165
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	61.36	27	17	44
	Yes	84.62	14	77	91
	Total	77.04	41	92	135
C.	Logit Regression (Mixed Effects) (LRME)				
	No	65.91	29	15	44
	Yes	92.31	7	84	91
	Total	83.70	36	99	135

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	58.82	20	14	34
	Yes	81.82	16	72	88
	Total	75,41	30	86	122

Table A-8.

Holdout Analysis Number 5

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	79.59	39	10	49
	Yes	87.86	17	123	140
	Total	85.71	56	133	189
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	82.22	37	8	45
	Yes	97.39	3	112	115
	Total	93.13	40	120	160
C.	Logit Regression (Mixed Effects) (LRME)				
	No	80.00	36	9	45
	Yes	88.70	13	102	115
	Total	86.25	49	111	160
D.	LDA* (adjusted for cases and effects)				
	No	81.40	35	8	43
	Yes	96.08	4	98	102
	Total	91.72	39	106	145

Table A-8.1

Holdout Analysis Number 5

	Observed	% correct	Classified No	Classified Yes	Total
A.	Linear Discriminant Analysis (LDA)				
	No	76.32	29	9	38
	Yes	80.36	22	90	112
	Total	79.33	51	99	150
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	61.76	21	13	34
	Yes	85.39	13	76	89
	Total	78.86	34	89	123
C.	Logit Regression (Mixed Effects) (LRME)				
	No	73.53	25	9	34
	Yes	86.52	12	77	89
	Total	82.93	37	86	123

	Observed	% correct	Classified No	Classified Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	72.22	26	10	36
	Yes	76.47	24	78	102
	Total	75.36	50	88	138

Table A-8

Holdout Analysis Number 6

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	85.29	58	10	68
	Yes	84.85	30	168	198
	Total	84.96	88	178	266
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	92.06	58	5	63
	Yes	94.34	9	150	159
	Total	93.69	67	155	222
C.	Logit Regression (Mixed Effects) (LRME)				
	No	80.95	51	12	63
	Yes	94.97	8	151	159
	Total	90.99	59	163	222

D.	LDA* (adjusted for cases and effects)				
	No	89.06	57	7	64
	Yes	90.12	16	146	162
	Total	89.82	73	153	226

Table A-8.1

Holdout Analysis Number 6

Observed	% correct	Classified No	Classified Yes	Classified Total
A: Linear Discriminant Analysis (LDA)				
No	73.68	14	5	19
Yes	79.63	11	43	54
Total	78.08	25	48	73
B. Logit Regression (Fixed Effects) (LRFE)				
No	68.75	11	5	16
Yes	77.78	10	35	45
Total	75.41	21	40	61
C. Logit Regression (Mixed Effects) (LRME)				
No	56.25	9	7	16
Yes	95.56	2	43	45
Total	85.25	11	50	61

Observed	% correct	No	Yes	Total
D. LDA* (adjusted for cases and effects)				
No	60.00	9	6	15
Yes	80.95	8	34	42
Total	75.44	17	40	57

Table A-8

Holdout Analysis Number 7

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	77.14	54	16	70
	Yes	88.78	22	174	196
	Total	85.71	76	190	266
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	80.95	51	12	63
	Yes	94.94	8	150	158
	Total	90.95	59	162	221
C.	Logit Regression (Mixed Effects) (LRME)				
	No	82.54	52	11	63
	Yes	89.87	16	142	158
	Total	87.78	68	153	221

D.	LDA* (adjusted for cases and effects)				
	No	83.33	50	10	60
	Yes	93.33	11	154	165
	Total	90.66	61	164	225

Table A-8.1

Holdout Analysis Number 7

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	70.59	12	5	17
	Yes	87.50	7	49	56
	Total	83.56	19	54	73
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	62.50	10	6	16
	Yes	93.48	3	43	46
	Total	85.48	13	49	62
C.	Logit Regression (Mixed Effects) (LRME)				
	No	62.50	10	6	16
	Yes	95.65	2	44	46
	Total	87.10	12	50	62
D.	LDA* (adjusted for cases and effects)				
	No	57.89	11	8	19
	Yes	79.49	8	31	39
	Total	72.41	19	39	58

Table A-8

Holdout Analysis Number 8

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	74.67	56	19	75
	Yes	93.43	13	185	198
	Total	88.28	69	204	273
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	89.55	60	7	67
	Yes	95.09	8	155	163
	Total	93.48	68	162	230
C.	Logit Regression (Mixed Effects) (LRME)				
	No	80.60	54	13	67
	Yes	93.87	10	153	163
	Total	90.00	64	166	230

D.	LDA* (adjusted for cases and effects)				
	No	85.29	58	10	68
	Yes	90.06	17	154	171
	Total	88.70	75	164	239

Table A-8.1

Holdout Analysis Number 8

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	25.00	3	9	12
Yes	92.59	4	50	54
Total	80.30	7	59	66
B. Logit Regression (Fixed Effects) (LRFE)				
No	41.67	5	7	12
Yes	87.80	5	36	41
Total	77.36	10	43	53
C. Logit Regression (Mixed Effects) (LRME)				
No	33.33	4	8	12
Yes	90.24	4	37	41
Total	77.36	8	45	53

Observed	% correct	Classified No	Classified Yes	Classified Total
D. LDA* (adjusted for cases and effects)				
No	45.45	5	6	11
Yes	84.85	5	28	33
Total	75.00	10	34	44

Table A-8

Holdout Analysis Number 9

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	75.71	53	17	70
Yes	90.05	19	172	191
Total	86.21	72	189	261
B. Logit Regression (Fixed Effects) (LRFE)				
No	95.38	62	3	65
Yes	89.81	16	141	157
Total	91.44	78	144	222
C. Logit Regression (Mixed Effects) (LRME)				
No	84.62	55	10	65
Yes	89.81	16	141	157
Total	88.29	71	151	222

	% correct	No	Yes	Total
D. LDA* (adjusted for cases and effects)				
No	79.69	51	13	64
Yes	93.29	11	153	164
Total	89.47	62	166	228

Table A-8.1

Holdout Analysis Number 9

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	64.71	11	6	17
	Yes	88.52	7	54	61
	Total	83.33	18	60	78
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	78.57	11	3	14
	Yes	76.60	11	36	47
	Total	77.05	22	39	61
C.	Logit Regression (Mixed Effects) (LRME)				
	No	64.29	9	5	14
	Yes	91.49	4	43	47
	Total	85.25	13	48	61
D.	LDA* (adjusted for cases and effects)				
	No	40.00	6	9	15
	Yes	87.5	5	35	40
	Total	74.55	11	44	55

Table A-8

Holdout Analysis Number 10

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	85.92	61	10	71
Yes	87.50	26	182	208
Total	87.10	87	192	279
B. Logit Regression (Fixed Effects) (LRFE)				
No	90.63	58	6	64
Yes	92.81	12	155	167
Total	92.21	70	161	231
C. Logit Regression (Mixed Effects) (LRME)				
No	87.50	56	8	64
Yes	90.42	16	151	167
Total	89.61	72	159	231

Observed	% correct	Classified No	Classified Yes	Classified Total
D. LDA* (adjusted for cases and effects)				
No	79.03	49	13	62
Yes	90.51	15	143	158
Total	87.27	79	141	220

Table A-8.1

Holdout Analysis Number 10

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	75.00	12	4	16
	Yes	79.55	9	35	44
	Total	78.33	21	39	60
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	80.00	12	3	15
	Yes	72.97	10	27	37
	Total	75.00	22	30	52
C.	Logit Regression (Mixed Effects) (LRME)				
	No	60.00	9	6	15
	Yes	97.30	1	36	37
	Total	86.54	10	42	52
D.	LDA* (adjusted for cases and effects)				
	No	76.47	13	4	17
	Yes	86.96	6	40	46
	Total	84.13	19	44	63

TABLE A-9

Table A-9

Bootstrap Analysis Number 1

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	79.31	69	18	87
	Yes	94.44	14	238	252
	Total	90.56	83	356	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	92.86	77	6	83
	Yes	93.90	13	200	213
	Total	93.60	90	206	296
C.	Logit Regression (Mixed Effects) (LRME)				
	No	90.36	75	8	83
	Yes	92.96	15	198	213
	Total	92.23	90	206	296
D.	LDA* (adjusted for cases and effects)				
	No	80.90	72	17	89
	Yes	93.81	12	182	194
	Total	89.75	84	199	283

Table A-9.1

Bootstrap Analysis Number 1

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	56.32	49	38	87
	Yes	92.46	19	233	252
	Total	83.19	68	271	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	65.82	52	27	79
	Yes	92.16	16	188	204
	Total	84.81	68	215	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	74.68	59	20	79
	Yes	88.73	23	181	204
	Total	84.81	82	201	283
D.	LDA* (adjusted for cases and effects)				
	No	73.42	58	21	79
	Yes	86.76	27	177	204
	Total	83.04	85	198	283

Table A-9

Bootstrap Analysis Number 2

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	75.51	74	24	98
	Yes	92.95	17	224	241
	Total	87.91	91	248	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	87.91	80	11	91
	Yes	97.57	5	201	206
	Total	94.61	85	212	297
C.	Logit Regression (Mixed Effects) (LRME)				
	No	90.11	82	9	91
	Yes	89.32	22	184	206
	Total	89.56	104	193	297

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	82.14	69	15	84
	Yes	87.94	24	175	199
	Total	86.22	93	290	283

Table A-9.1

Bootstrap Analysis Number 2

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	71.26	62	25	87
	Yes	90.87	23	229	252
	Total	85.84	85	254	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	78.48	62	17	79
	Yes	80.39	40	164	204
	Total	79.86	102	181	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	83.54	66	13	79
	Yes	80.88	39	165	204
	Total	81.63	105	178	283

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	70.89	56	23	79
	Yes	90.20	20	184	204
	Total	84.81	76	207	283

Table A-9

Bootstrap Analysis Number 3

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	75.29	64	21	85
	Yes	90.94	23	231	254
	Total	87.02	87	252	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	93.67	74	5	79
	Yes	92.86	15	195	210
	Total	93.08	89	200	289
C.	Logit Regression (Mixed Effects) (LRME)				
	No	86.08	68	11	79
	Yes	92.38	16	194	210
	Total	90.66	84	205	289

	Observed	% correct	No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	89.74	70	8	78
	Yes	90.24	20	185	205
	Total	90.11	90	193	283

Table A-9.1

Bootstrap Analysis Number 3

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	73.56	64	23	87
	Yes	88.10	30	222	252
	Total	84.37	94	245	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	79.75	63	16	79
	Yes	85.29	3-	174	204
	Total	83.75	93	190	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	73.42	58	21	79
	Yes	91.18	18	186	204
	Total	86.22	76	207	283

D.	LDA* (adjusted for cases and effects)				
	No	82.28	65	14	79
	Yes	87.25	26	178	204
	Total	85.87	91	192	283

Table A-9

Bootstrap Analysis Number 4

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	88.89	72	9	81
	Yes	89.92	26	232	258
	Total	89.68	98	241	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	83.12	64	13	77
	Yes	95.73	9	202	211
	Total	92.36	73	215	288
C.	Logit Regression (Mixed Effects) (LRME)				
	No	92.21	71	6	77
	Yes	86.73	28	183	211
	Total	88.19	99	189	288
D.	LDA* (adjusted for cases and effects)				
	No	90.72	88	9	97
	Yes	92.47	14	172	186
	Total	91.87	102	181	283

Table A-9.1

Bootstrap Analysis Number 4

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	77.01	67	20	87
	Yes	89.68	26	226	252
	Total	86.43	93	246	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	64.56	51	28	79
	Yes	93.63	13	191	204
	Total	85.51	64	219	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	86.08	68	11	79
	Yes	82.84	35	169	204
	Total	83.75	103	180	283

			No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	79.75	63	16	79
	Yes	84.80	31	173	204
	Total	83.39	94	189	283

Table A-9

Bootstrap Analysis Number 5

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	76.40	68	21	89
	Yes	95.20	12	238	250
	Total	90.27	80	259	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	87.65	71	10	81
	Yes	100.00	0	201	201
	Total	96.45	71	211	282
C.	Logit Regression (Mixed Effects) (LRME)				
	No	82.72	67	14	81
	Yes	96.52	7	194	201
	Total	92.55	74	208	282

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	84.27	75	14	89
	Yes	91.75	16	178	194
	Total	89.40	91	192	283

Table A-9.1

Bootstrap Analysis Number 5

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	66.67	58	29	87
	Yes	91.67	21	231	252
	Total	85.25	79	260	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	55.70	44	35	79
	Yes	93.14	14	190	204
	Total	82.69	58	225	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	74.68	59	20	79
	Yes	92.65	15	189	204
	Total	87.63	74	209	283
D.	LDA* (adjusted for cases and effects)				
	No	77.22	61	18	79
	Yes	86.76	27	177	204
	Total	84.10	88	195	283

Table A-9

Bootstrap Analysis Number 6.

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	73.96	71	25	96
	Yes	93.00	17	226	243
	Total	87.61	88	251	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	94.51	86	5	91
	Yes	93.33	13	182	195
	Total	93.71	99	187	286
C.	Logit Regression (Mixed Effects) (LRME)				
	No	94.51	86	5	91
	Yes	89.74	20	175	195
	Total	91.26	106	180	286
D.	LDA* (adjusted for cases and effects)				
	No	91.43	64	6	70
	Yes	91.55	18	195	213
	Total	91.52	82	201	283

Table A-9.1

Bootstrap Analysis Number 6

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	64.37	56	31	87
	Yes	93.65	16	236	252
	Total	86.14	72	267	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	78.48	62	17	79
	Yes	85.29	30	174	204
	Total	83.39	92	191	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	68.35	54	25	79
	Yes	92.16	16	188	204
	Total	85.51	70	213	283

D.	LDA* (adjusted for cases and effects)				
	No	81.01	64	15	79
	Yes	87.75	25	179	204
	Total	85.87	89	194	283

Table A-9

Bootstrap Analysis Number 7

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	85.87	79	13	92
	Yes	91.90	20	227	247
	Total	90.27	99	240	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	95.12	78	4	82
	Yes	99.02	2	203	205
	Total	97.91	80	207	287
C.	Logit Regression (Mixed Effects) (LRME)				
	No	90.24	74	8	82
	Yes	90.69	20	185	205
	Total	90.24	94	193	287

	Observed	% correct	No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	82.72	67	14	81
	Yes	94.55	11	191	202
	Total	91.17	78	205	283

Table A-9.1

Bootstrap Analysis Number 7

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	72.41	63	24	87
Yes	86.90	33	219	252
Total	83.19	96	243	339
B. Logit Regression (Fixed Effects) (LRFE)				
No	78.48	62	17	79
Yes	84.80	31	173	204
Total	83.04	93	190	283
C. Logit Regression (Mixed Effects) (LRME)				
No	67.09	53	26	79
Yes	91.67	17	187	204
Total	84.81	70	213	283
D. LDA* (adjusted for cases and effects)				
No	79.75	63	16	79
Yes	90.20	20	184	204
Total	87.28	83	200	283

Table A-9

Bootstrap Analysis Number 8

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	90.54	67	7	74
	Yes	87.55	33	232	265
	Total	88.20	100	239	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	91.43	64	6	70
	Yes	93.33	14	196	210
	Total	92.86	78	202	280
C.	Logit Regression (Mixed Effects) (LRME)				
	No	85.71	67	3	70
	Yes	89.52	22	188	210
	Total	91.07	89	191	280

D.	LDA* (adjusted for cases and effects)				
	No	79.31	69	18	87
	Yes	94.90	10	186	196
	Total	90.11	79	204	283

Table A-9.1

Bootstrap Analysis Number 8

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	73.56	64	23	87
	Yes	87.30	32	220	252
	Total	83.78	96	243	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	75.95	60	19	79
	Yes	89.22	22	182	204
	Total	85.51	82	201	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	79.75	63	16	79
	Yes	88.73	23	181	204
	Total	86.22	86	197	283

D.	LDA* (adjusted for cases and effects)				
	No	65.82	52	27	79
	Yes	91.67	17	187	204
	Total	84.45	69	214	283

Table A-9

Bootstrap Analysis Number 9

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	85.00	68	12	80
	Yes	89.96	26	233	259
	Total	88.79	94	245	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	93.42	71	5	76
	Yes	91.09	18	184	202
	Total	91.73	89	189	278
C.	Logit Regression (Mixed Effects) (LRME)				
	No	93.42	71	5	76
	Yes	88.12	24	178	202
	Total	89.57	95	183	278

D.	LDA* (adjusted for cases and effects)				
	No	80.28	57	14	71
	Yes	90.09	21	191	212
	Total	87.63	78	205	283

Table A-9.1

Bootstrap Analysis Number 9

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	67.82	59	28	87
	Yes	88.89	28	224	252
	Total	83.48	87	252	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	78.48	62	17	79
	Yes	81.86	37	167	204
	Total	80.92	99	284	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	86.08	68	11	79
	Yes	81.37	38	166	204
	Total	82.69	106	177	283

D.	LDA* (adjusted for cases and effects)				
	No	70.89	56	23	79
	Yes	91.18	18	186	204
	Total	85.51	74	209	283

Table A-9

Bootstrap Analysis Number 10

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	87.01	67	10	77
	Yes	91.60	22	240	262
	Total	90.56	89	250	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	91.43	64	6	70
	Yes	100.00	0	214	214
	Total	97.89	64	220	284
C.	Logit Regression (Mixed Effects) (LRME)				
	No	94.29	66	4	70
	Yes	91.59	18	196	214
	Total	92.25	84	200	284

D.	LDA* (adjusted for cases and effects)				
	No	81.58	62	14	76
	Yes	93.72	13	194	207
	Total	90.46	75	208	283

Table A-9.1

Bootstrap Analysis Number 10

	Observed	% correct	No	Yes	Total
			Classified		
A.	Linear Discriminant Analysis (LDA)				
	No	72.41	63	24	87
	Yes	89.68	26	226	252
	Total	85.25	89	250	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	79.75	63	16	79
	Yes	63.89	43	161	252
	Total	79.15	106	177	339
C.	Logit Regression (Mixed Effects) (LRME)				
	No	82.28	65	14	79
	Yes	87.75	25	179	204
	Total	86.22	90	193	283
D.	LDA* (adjusted for cases and effects)				
	No	75.95	60	19	79
	Yes	89.71	21	183	204
	Total	85.87	81	202	283

Table A-9

Bootstrap Analysis Number 11

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	91.95	80	7	87
	Yes	89.68	26	226	252
	Total	90.27	106	233	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	96.00	72	3	75
	Yes	97.60	5	203	208
	Total	97.17	77	206	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	84.00	63	12	75
	Yes	95.19	10	198	208
	Total	92.23	73	210	283

D.	LDA* (adjusted for cases and effects)				
	No	84.21	64	12	76
	Yes	94.20	12	195	207
	Total	91.52	76	207	283

Table A-9.1

Bootstrap Analysis Number 11

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	81.61	71	16	87
	Yes	86.11	35	217	252
	Total	84.96	106	233	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	79.75	63	16	79
	Yes	85.78	29	175	204
	Total	84.10	92	191	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	75.95	60	19	79
	Yes	90.20	20	184	204
	Total	86.22	80	203	283
D.	LDA* (adjusted for cases and effects)				
	No	77.22	61	18	79
	Yes	87.25	26	178	204
	Total	84.45	87	196	283

Table A-9

Bootstrap Analysis Number 12

	Observed	% correct	No	Classified Yes	Total
A.	Linear Discriminant Analysis (LDA)				
	No	88.89	64	8	72
	Yes	93.26	18	249	267
	Total	92.33	82	257	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	95.45	63	3	66
	Yes	99.01	2	201	203
	Total	98.14	65	204	269
C.	Logit Regression (Mixed Effects) (LRME)				
	No	87.88	58	8	66
	Yes	96.06	8	195	203
	Total	94.05	66	203	269

D.	LDA* (adjusted for cases and effects)				
	No	90.70	78	8	86
	Yes	94.92	10	187	197
	Total	93.64	88	195	283

Table A-9.1

Bootstrap Analysis Number 12

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	54.02	47	40	87
	Yes	96.83	8	244	252
	Total	85.84	55	284	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	70.89	56	23	79
	Yes	91.67	17	187	204
	Total	85.87	73	210	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	73.42	58	21	79
	Yes	90.69	19	185	204
	Total	85.87	77	206	283

D.	LDA* (adjusted for cases and effects)				
	No	79.75	63	16	79
	Yes	87.25	26	178	204
	Total	85.16	89	194	283

Table A-9

Bootstrap Analysis Number 13

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	87.10	81	12	93
	Yes	91.87	20	226	246
	Total	90.56	101	238	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	91.46	75	7	82
	Yes	100.00	0	193	193
	Total	97.45	75	200	275
C.	Logit Regression (Mixed Effects) (LRME)				
	No	89.02	73	9	82
	Yes	94.30	11	182	193
	Total	92.73	84	191	275

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	79.76	67	17	84
	Yes	94.97	10	189	199
	Total	90.46	77	206	283

Table A-9.1

Bootstrap Analysis Number 13

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	62.07	54	33	87
	Yes	90.87	23	229	252
	Total	83.48	77	262	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	75.95	60	19	79
	Yes	90.20	20	184	204
	Total	86.22	80	203	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	81.01	64	15	79
	Yes	85.78	29	175	204
	Total	84.45	93	190	283

D.	LDA* (adjusted for cases and effects)				
	No	70.89	56	23	79
	Yes	90.20	20	184	204
	Total	84.81	76	207	283

Table A-9.

Bootstrap Analysis Number 14

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	84.15	69	13	82
	Yes	97.94	31	226	257
	Total	87.02	100	239	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	84.00	63	12	75
	Yes	97.18	6	207	213
	Total	93.75	69	219	288
C.	Logit Regression (Mixed Effects) (LRME)				
	No	80.00	60	15	75
	Yes	93.90	13	200	213
	Total	90.28	73	215	288
D.	LDA* (adjusted for cases and effects)				
	No	91.78	67	6	73
	Yes	98.10	4	206	210
	Total	96.47	71	212	283

Table A-9.1

Bootstrap Analysis Number 14

	Observed	% correct	Classified No	Classified Yes	Total
A.	Linear Discriminant Analysis (LDA)				
	No	72.41	63	24	87
	Yes	87.30	32	220	252
	Total	83.48	95	244	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	69.62	55	24	79
	Yes	94.12	12	192	204
	Total	87.28	67	216	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	75.95	60	19	79
	Yes	88.73	23	181	204
	Total	85.16	83	200	283

D.	LDA* (adjusted for cases and effects)				
	No	69.62	55	24	79
	Yes	90.69	19	185	204
	Total	84.81	74	209	283

Table A-9

Bootstrap Analysis Number 15

	Observed	% Correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	74.07	60	21	81
	Yes	96.90	8	250	258
	Total	91.45	68	271	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	92.21	71	6	77
	Yes	95.22	10	199	209
	Total	94.41	81	205	286
C.	Logit Regression (Mixed Effects) (LRME)				
	No	84.42	65	12	77
	Yes	94.74	11	198	209
	Total	91.96	76	210	286

D.	LDA* (adjusted for cases and effects)				
	No	96.25	77	3	80
	Yes	91.63	17	186	203
	Total	92.93	94	189	283

Table A-9.1

Bootstrap Analysis Number 15

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	52.87	46	41	87
	Yes	96.03	10	242	252
	Total	84.96	56	283	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	70.89	56	23	79
	Yes	90.69	19	185	204
	Total	85.16	75	208	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	67.09	53	26	79
	Yes	93.14	14	190	204
	Total	85.87	67	216	283

D.	LDA* (adjusted for cases and effects)				
	No	87.34	69	10	79
	Yes	82.35	36	168	204
	Total	83.75	105	178	283

Table A-9.

Bootstrap Analysis Number 16

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	82.14	69	15	84
	Yes	89.41	27	228	255
	Total	87.61	96	243	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	79.22	61	16	77
	Yes	97.51	5	196	201
	Total	92.45	66	212	278
C.	Logit Regression (Mixed Effects) (LRME)				
	No	93.51	72	5	77
	Yes	86.07	28	173	201
	Total	88.13	100	178	278

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	88.73	63	8	71
	Yes	94.34	12	200	212
	Total	92.93	75	208	283

Table A-9.1

Bootstrap Analysis Number 16

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	80.46	70	17	87
	Yes	85.71	36	216	252
	Total	84.37	106	233	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	67.09	53	26	79
	Yes	94.12	12	192	204
	Total	86.57	65	218	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	86.08	68	11	79
	Yes	79.41	42	162	204
	Total	81.27	110	173	283

	Observed	% correct	No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	73.42	58	21	79
	Yes	87.75	25	179	204
	Total	83.75	83	200	283

Table A-9

Bootstrap Analysis Number 17

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	78.38	58	16	74
	Yes	90.57	25	240	265
	Total	87.91	83	256	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	86.57	58	9	67
	Yes	98.15	4	212	216
	Total	95.41	62	221	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	74.63	50	17	67
	Yes	96.30	8	208	216
	Total	91.17	58	225	283

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	81.58	62	14	76
	Yes	89.86	21	186	207
	Total	87.63	83	200	283

Table A-9.1

Bootstrap Analysis Number 17

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	64.37	56	31	87
	Yes	93.25	17	235	252
	Total	85.84	73	266	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	63.29	50	29	79
	Yes	87.75	25	179	204
	Total	80.92	75	208	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	58.23	46	33	79
	Yes	95.59	9	195	204
	Total	85.16	55	228	283

	Observed	% correct	No	Yes	Total
D.	LDA* (adjusted for cases and effects)				
	No	75.95	60	19	79
	Yes	90.69	19	185	204
	Total	86.57	79	204	283

Table A-9

Bootstrap Analysis Number 18

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	77.78	77	22	99
Yes	91.67	20	220	240
Total	87.61	97	242	339
B. Logit Regression (Fixed Effects) (LRFE)				
No	98.95	94	1	95
Yes	92.39	15	182	197
Total	94.54	109	183	292
C. Logit Regression (Mixed Effects) (LRME)				
No	96.91	94	3	97
Yes	90.26	19	176	195
Total	92.47	113	179	292

Observed	% correct	No	Yes	Total
D. LDA* (adjusted for cases and effects)				
No	79.17	57	15	72
Yes	93.36	14	197	211
Total	89.75	71	212	283

Table A-9.1

Bootstrap Analysis Number 18

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	73.56	64	23	87
	Yes	89.68	26	226	252
	Total	85.55	90	249	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	88.61	70	9	79
	Yes	75.49	50	154	204
	Total	79.15	120	163	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	82.28	65	14	79
	Yes	87.25	26	178	204
	Total	85.87	91	292	283
D.	LDA* (adjusted for cases and effects)				
	No	67.09	53	26	79
	Yes	91.17	17	187	204
	Total	84.81	70	213	283

Table A-9

Bootstrap Analysis Number 19

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	71.05	54	22	76
	Yes	93.16	18	245	263
	Total	88.20	72	267	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	94.37	67	4	71
	Yes	93.49	14	201	215
	Total	93.71	81	205	286
C.	Logit Regression (Mixed Effects) (LRME)				
	No	87.32	62	9	71
	Yes	92.56	16	199	215
	Total	91.26	78	208	286

D.	LDA* (adjusted for cases and effects)				
	No	81.01	64	15	79
	Yes	92.16	16	188	204
	Total	89.05	80	203	283

Table A-9.1

Bootstrap Analysis Number 19

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	60.92	53	34	87
	Yes	94.44	14	238	252
	Total	85.84	67	272	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	81.01	64	15	79
	Yes	88.73	23	181	204
	Total	86.57	87	196	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	84.81	67	12	79
	Yes	88.73	23	181	204
	Total	87.63	90	193	283

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA (adjusted for cases and effects)				
	No	74.68	59	20	79
	Yes	91.18	18	186	204
	Total	86.57	77	206	283

Table A-9

Bootstrap Analysis Number 20

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	88.89	80	10	90
	Yes	87.95	30	219	249
	Total	88.20	110	229	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	94.05	79	5	84
	Yes	92.16	16	188	204
	Total	92.71	95	193	288
C.	Logit Regression (Mixed Effects) (LRME)				
	No	76.19	64	20	84
	Yes	97.06	6	198	204
	Total	90.97	70	218	288

	Observed	% correct	Classified No	Classified Yes	Classified Total
D.	LDA* (adjusted for cases and effects)				
	No	77.65	66	19	85
	Yes	94.95	10	188	198
	Total	89.75	76	207	283

Table A-9.1

Bootstrap Analysis Number 20

	Observed	% correct	No	Yes	Total
			Classified		
A.	Linear Discriminant Analysis (LDA)				
	No	74.71	65	22	87
	Yes	91.27	22	230	252
	Total	87.02	87	252	339
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	77.22	61	18	79
	Yes	85.29	30	174	204
	Total	83.04	91	192	283
C.	Logit Regression (Mixed Effects) (LRME)				
	No	64.56	51	28	79
	Yes	93.63	13	191	204
	Total	85.51	64	219	283

D.	LDA* (adjusted for cases and effects)				
	No	72.15	57	22	79
	Yes	88.73	23	181	204
	Total	84.09	80	203	283

Table A-10

Missing Data: Models Fit on Complete Data Set

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	87.36	76	11	87
Yes	91.74	20	222	242
Total	90.58	96	233	329
% correct		79.17	95.28	
Logit Regression (Mixed Effects) (LRME)				
No	80.46	70	17	87
Yes	90.08	24	218	242
Total	87.54	94	235	329
% correct		74.47	92.77	

Table A-10.1
Missing Data: Original model used to predict new data set of predicted missing values

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	79.31	69	18	87
Yes	89.68	26	226	252
Total	87.02	95	244	339
% correct		72.63	92.62	
Logit Regression (Mixed Effects) (LRME)				
No	72.41	63	24	87
Yes	92.46	19	233	252
Total	87.32	82	257	339
% correct		76.83	90.66	

Table A-11

"Separation" Models - Classification

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	64.56	51	28	79
Yes	95.10	10	194	204
Total	86.57	61	222	283
% correct		83.61	87.39	
Logit Regression (Mixed Effects) (LRME)				
No	65.82	52	27	79
Yes	94.61	11	193	204
Total	86.57	63	220	283
% correct		82.54	87.73	
LRFE* (Model with variables forced in)				
No	68.35	54	25	79
Yes	94.61	11	193	204
Total	87.28	65	218	283
% correct		83.08	88.53	

TABLE A-12

COEFFICIENTS FOR THE "SEPARATION" MODELS

		LRFE	LRFE*	LRME
Constant		1.874	4.1722	2.5507
TY	(1)	0	0	0
	(2)	-.477	.2435	-.0271
	(3)	.233	-.1724	.6654
	(4)	.644	.6434	1.1264
NEW	(1)	0	0	0
	(2)	-1.401	-1.4879	-1.1302
	(3)	1.293	1.3514	1.3586
	(4)	.283	.1917	.4443
AMT		-.00063	-.00148	-.000449
AMT/TNW		.00037	.00201	.000189
U	(1)	0	0	0
	(2)	-1.3983	-3.2704	-2.7949
$BRN_{j(1)}$	(2)	1.9216	2.1149	1.1893
	(4)	-1.0953	-1.1438	-.7292
	(11)	.1056	.0204	.1102
	(19)	0	0	-.5703
$BRN_{j(2)}$	(1)	-	-3.4584	.0106
	(5)	-	4.6740	-.0623
	(6)	-	-3.6844	.0143
	(10)	-	∞	.0298
	(12)	-	∞	.0013
	(14)	-	∞	.0092
	(15)	-	∞	.0492
	(17)	-	∞	.0066
	(18)	-	∞	.0463
	(21)	-	-3.[illegible]67	-.0125
	(22)	-	∞	.0179
	(24)	-	∞	.0759
	(26)	-	∞	.0318
	(27)	-	-4.5809	-.0391
	(28)	-	∞	.0192
	(30)	-	-6.3819	-.0622
	(31)	-	∞	.0385
	(32)	-	∞	.0039
	(33)	-	-6.0611	-.0655
	(34)	-	-6.0992	-.0469
	(35)	-	2 ∞	.0106
	(36)	-	4.0779	.0482
	(37)	-	-4.8452	-.0429
	(38)	-	-∞	-.0644
	(39)	-	[illegible]	.0157
	(40)	-	-4.4195	-.0181
	(41)	-	∞	.0348
	(42)	-	-5.2986	.0335

		LRFE	LRFE*	LRME
$BRN_{j(2)}$	(43)	-	∞	.0223
	(44)	-	∞	.0161
	(45)	-	∞	.0059
	(46)	-	-5.2935	-.0528
	(47)	-	-6.2931	-.0993
	(48)	-	∞	.0052

where U (1) = remaining 44 branches
(2) = oversampled 4 branches

$BRN_{j(1)}$ = specific branches within oversampled group

$BRN_{j(2)}$ = specific branches within remaining 44 branches group

LRFE* = is as defined in section 5.8

TABLE A-13

POINTS ON GRID

$\sigma_1^2(x)$ \ $\sigma_2^2(y)$	0.000	.030	.050	.0811	.110	.150	.250	.300	.500	1.000
0.000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.250	.0000	.0000	.0000	.0000	.0079	.0842	.1379	.1011	.0003	.0000
0.500	.0000	.0000	.00008	.5622	.7368	.2983	.1139	.0777	.0007	.0000
0.850	.0000	.0021	.9187	.9513	.8962	.6913	.4172	.2019	.0319	.0000
1.044	.0000	.0944	.9752	1.000	.9423	.7743	.4703	.2212	.0732	.0000
1.200	.0000	.1062	.8911	.9705	.9615	.8429	.5731	.5100	.1129	.0000
2.200	.0000	.1156	.5665	.5986	.6155	.8399	.7522	.5428	.2731	.0000
3.000	.0000	.0836	.2654	.4088	.4921	.6542	.6037	.5933	.3091	.0000
4.000	.0000	.0592	.1989	.3641	.3735	.3927	.4889	.4426	.3636	.0000
7.000	.0000	.0064	.0092	.0364	.0447	.0991	.1257	.0628	.0742	.0000
10.000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.00000

THREE-DIMENSIONAL PLOT OF LIKELIHOOD FUNCTION

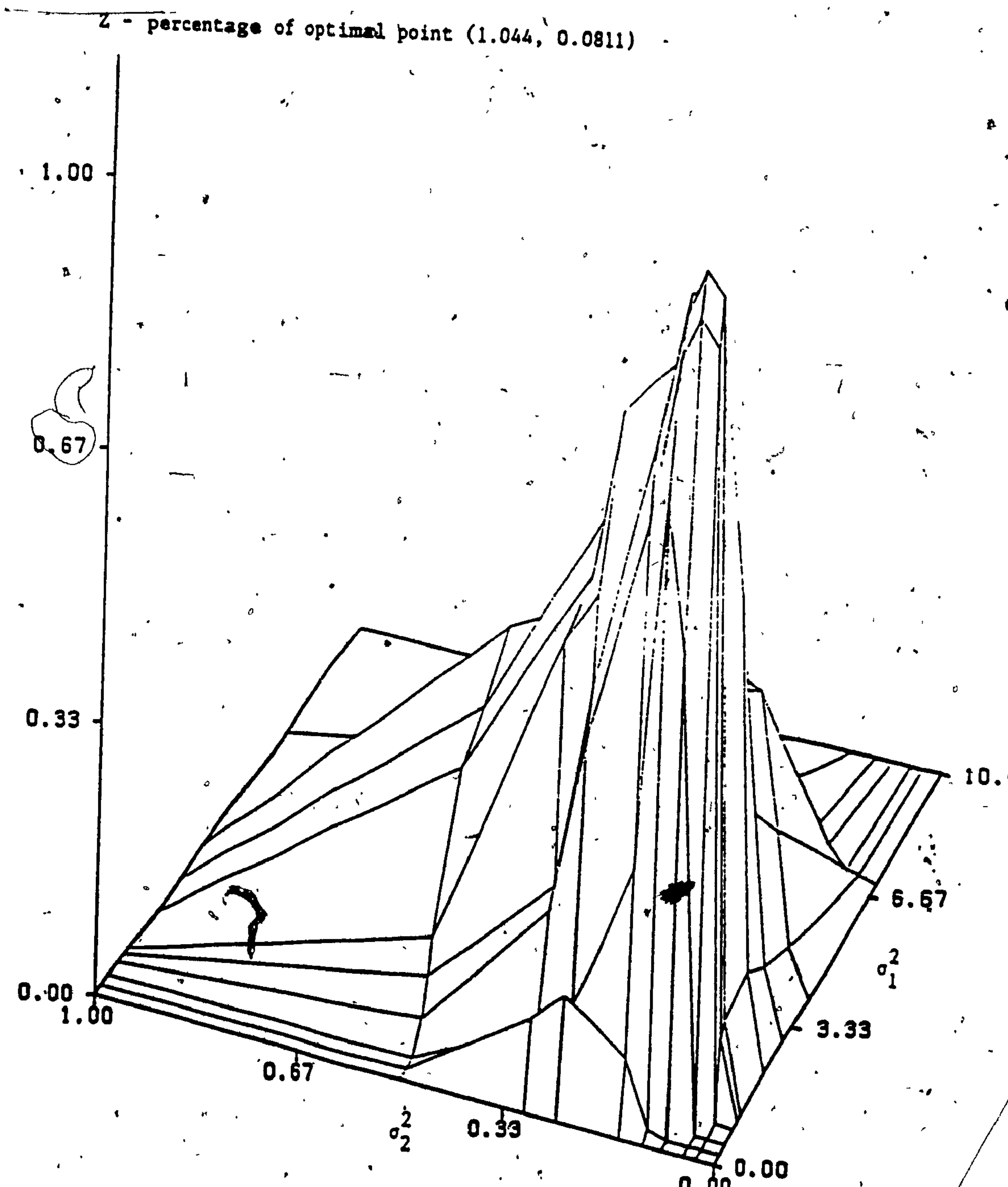


Table A -14
Division Management Level Classification

Observed	% correct	Classified No	Classified Yes	Classified Total
A. Linear Discriminant Analysis (LDA)				
No	30.77	8	18	26
Yes	98.81	3	249	252
Total	92.45	11	267	278
% correct		72.73	93.26	
B. Logit Regression (Fixed Effects) (LRFE)				
No	58.33	14	10	24
Yes	98.53	3	201	204
Total	94.30	17	211	228
% correct		82.35	95.26	
C. Logit Regression (Mixed Effects) (LRME)				
No	58.33	14	10	24
Yes	91.18	18	186	204
Total	87.72	32	196	228
% correct		43.75	94.90	

TABLE A - 15

Coefficients for the Division Management Level

	LDA	LRFE	LRME
Constant	.72134	8.1118	3.592
TY(1)		0	0
(2)		-.3783	-1.0481
(3)		-3.8687	.0175
(4)		.29724	-.1432
NEW(1)	0	0	0
(2)	.4560	-3.9002	-.9497
(3)	0	14.243	-.0861
(4)	0	-1.3946	.0401
AMT		-.005481	-.00229
AMTTN		-.001405	.000133
BRN 1	.000	0	.2385
2	.000	.365[illegible]	.0775
4	-0.9712	.7821	.2300
5	.067	1.9840	.1498
6	.003	-2.3695	-.00079
10	-.092	2. ∞	.2238
11	1.5009	-3.6[illegible]	-.9814
12	.000	∞	.0235
14	.000	∞	.0695
15	.000	∞	.3885
17	.004	∞	.0795
18	.539	∞	.2502
19	.052	32.4	.1077
21	.000	-2.2103	-.0489
22	.007	∞	.1105
24	.000	∞	.6181
26	0.000	∞	.1259
27	0.6510	-.7830	.01101
28	-0.1700	∞	.1245
30	-2.8409	-6.3634	-.3071
31	.001	∞	.4569
32	-.001	∞	.0243
33	.007	44.464	-.5166
34	9.2106	-19.902	-.3726
35	.051	∞	.1522
36	-.006	9.7253	.2979
37	2.3912	-34.377	-.4701
38	7.0192	-∞	-.6600
39	.071	∞	.0556
40	.000	-.02007	-.1053
41	.000	∞	.2329
42	-2.6863	58.902	-.2140
43	.005	∞	.1581
44	.007	∞	.1221
45	.001	∞	.0448

		LDA	LRFE	LRME
BRN	46	.000	363.91	-.2635
	47	.000	-2.2326	-.4685
	48	0	∞	.0451
BRN*TNW	1		0	-.0001
	2		.0005	.0004
	4		.0529	.0001
	5		.0417	.0001
	6		.0086	.0006
	10		.0349	.0001
	11		.0108	.0015
	12		.1009	.0000
	14		.0344	.0000
	15		-.0104	.0001
	17		-.5436	.0000
	18		-.00707	.0002
	19		-.1414	.0001
	21		.0049	.0008
	22		.0085	.0001
	24		.0097	.0005
	26		.0043	.0001
	27		-.0021	-.0002
	28		.00605	.0004
	30		.00807	-.0001
	31		.0060	.0003
	32		.0099	.0000
	33		-.7756	-.0002
	34		.2859	.0000
	35		-.0074	.0000
	36		-.0194	.0000
	37		.7561	.0000
	38		-1.4586	-.0001
	39		.0709	.0000
	40		.3911	.0002
	41		-.6971	.0001
	42		-1.364	-.0001
	43		.0470	.0000
	44		.1263	.0000
	45		.0246	.0000
	46		-3.4032	-.0002
	47		-.00491	-.0011
	48		-.3753	.0000
BRN*AMT	1	.00134		
	6	.00377		
	27	.00515		
	30	.01900		
	33	.01959		
	34	-.08766		
	42	.04260		
	46	.00851		
	47	.01705		

Table A-16

Branch Management Level Classification

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	96.72	59	2	61
	Yes	85.61	40	238	278
	Total	87.61	99	240	339
	% correct		59.60	99.17	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	86.54	45	7	52
	Yes	96.88	6	186	192
	Total	94.67	51	193	244
	% correct		88.24	96.37	

Table A-16

Branch Management Level Without BRN.

Observed	% correct	Classified No	Classified Yes	Total
A. Linear Discriminant Analysis (LDA)				
No	83.33	45	9	54
Yes	56.44	98	127	225
Total	61.65	143	136	279
% correct		31.47	93.38	
B. Logit Regression (Fixed Effects) (LRFE)				
No	51.92	27	25	52
Yes	93.75	12	180	192
Total	84.84	39	205	244
% correct		69.23	87.80	

TABLE A - 17

Coefficients for Branch Management Level

		LDA	LRFE
Constant		1.77919	-11.414
TY	1	-.16611	0
	2	0	-10.089
	3	0	.30915
	4	0	-1.7746
NEW	1	-.29352	0
	2	0	2.5643
	3	0	.70215
	4	0	-.22422
AMTTN			-.03300
BRN	1	.20974	0
	2	-1.07855	8.3068
	4	-3.4707	13.195
	5	-.44911	-1.1928
	6	-.32322	.63068
	10	-1.0058	∞
	11	-2.8849	10.[illegible]
	12	-.92781	∞
	14	-.37665	∞
	15	.1024	∞
	17	-.62869	∞
	18	-.37811	∞
	19	-3.9314	13.832
	21	-.21344	.58144
	22	1.0424	∞
	24	-.41573	∞
	26	.11139	∞
	27	-.60398	2.2406
	28	1.04236	∞
	30	-.37265	2.2728
	31	-.45154	∞
	32	-.45079	∞
	33	.23582	.06561
	34	-.51350	9.372[illegible]
	35	-.8300	∞
	36	-.30728	-.74049
	37	.47308	-.58437
	39	.09621	∞
	40	1.04236	.88700
	41	1.04236	∞
	42	-.50651	2.4275
	43	-.35701	∞
	44	-.54956	∞
	45	-.73368	∞
	46	-.58385	1.1872

	LDA	LRFE
BRN 47	-.66766	9.3801
48	0	∞
TYNEW (1,1)	0	
(1,2)	-.35811	
(1,3)	0	
(1,4)	-.03510	
(2,1)	-.00228	
(2,2)	-1.70853	
(2,3)	-.61775	
(2,4)	-.42875	
(3,1)	1.07208	
(3,2)	.28735	
(3,3)	-.6169	
(3,4)	-.02315	
(4,1)	-.09545	
(4,2)	-.10092	
(4.3)	-1.87501	
(4,4)	0	
BRN*AMT 1	-.00055	
2	.00062	
3	-.04479	
4	.00017	
5	.00426	
6	.00051	
7	0	
8	-.00828	
9	-.02289	
10	.00519	
11	.00219	
12	.03744	

Coefficients for Branch Management Level without BRN

	LDA	LRFE
Constant	1.43987	0.24961
TY 1	-2.00149	0
2	0	-10.100
3	0	-1.9412
4	0	-2.5206
NEW 1	-.89461	0
2	-1.3095	.03796
3	0	-.88481
4	0	-2.0277
TNW	.00032	-.00107
AMTTN	—	-.00446
AMT	.00045	—

APPENDIX B

```
      PROGRAM 1                                                    NLIN
(INPUT,OUTPUT,TAPE23=INPUT,TAPE6=OUTPUT,TAPE4,TAPE7)
      COMMON /VALU/ Y,X1,X2,X3,SS1,SS2
      DIMENSION Y(34,4,4,31) , X1(34,4,4,31), X2(34,4,4,31),
     1    X3(34,4,4,31)
      DIMENSION X(77), F(77), PAR(77), WK(12000)
      EXTERNAL FCN
      DATA N,NSIG,ITMAX/77,5,600/
      READ (4,*) X,S1,S2
      TEST
       PRINT*,X
      READ (7,*) IDGT
      READ (7,*) A,AINV
      READ (7,*) A,AINV
      READ (7,*) A,AINV
      SS1 = 0.
      SS2 = 0.
      DO 5 I=1,34
      READ (7,*) A,AINV
      SS1 = SS1 + (X(I+3)**2) + AINV
  5    CONTINUE
      DO 6 I=1,6
      READ (7,*) A,AINV
  6    CONTINUE
      DO 7 I=1,34
      READ (7,*) A,AINV
      SS2 = SS2 + (X(I+43)**2) + AINV
  7    CONTINUE
      SS1 = SS1/34
      SS2 = SS2/34
C       PRINT*,SS1,SS2
C     SS1 = 1.8928
C     SS2 = .0000041612252
      IF (((SS1.GE.(0.99*S1)).AND.(SS1.LE.(1.01*S1)))  .AND.
      TEST
     1    ((SS2.GE.(0.99*S2)).AND.(SS2.LE.(1.01*S2)))) CALL DUMMY
     TEST
       CALL READ
      CALL ZSPOW (FCN,NSIG,N,ITMAX,PAR,X,FNORM,WK,IER)
      STOP
      END
      SUBROUTINE READ
      COMMON /VALU/ Y,X1,X2,X3,SS1,SS2
      DIMENSION Y(34,4,4,31),X1(34,4,4,31),X2(34,4,4,31),
     1    X3(34,4,4,31)
      II = 1
      III = 1
 10   READ (23,15,END=100) A1,A2,K,I,J,A3,L
 15   FORMAT (4X,F1.0,F4.0,5X,I1,I2,5X,I1,27X,F5.0,16X,I2)
      IF (A3 .EQ. 99999) GO TO 10
      IF (III.NE.I) II = II + 1
      Y(II,J,K,L) = A1
      X1(II,J,K,L) = A2
      X3(II,J,K,L) = A3
```

```
      X2(II,J,K,L) = X9(II,J,K,L) /(X3(II,J,K,L))
      III = I
      GO TO 10
100   RETURN
      END
      SUBROUTINE FCN (X,F,N,PAR)
      COMMON /VALU/ Y,X1,X2,X3,SS1,SS2
      DIMENSION Y(34,4,4,31), X1(34,4,4,31), X2(34,4,4,31),
     1      X3(34,4,4,31)
      DIMENSION X(N),F(N),PAR(1)
      PRINT*,X
      REWIND 1
      WRITE (1,*) X, SS1, SS2
      YY = 0.
      PP = 0.
      DO 10 I=1,31
      DO 10 J=1,4
      DO 10 K=1,4
      DO 10 L=1,31
      IF ( X1(I,J,K,L) .EQ. 0.0 ) GO TO 10
      YY = YY + Y(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)
      PP = PP + WW
10    CONTINUE
      F(1) = YY - PP
       PRINT*,YY,PP
      YY = 0.
      PP = 0.
      DO 20 I=1,34
      DO 20 J=1,4
      DO 20 K=1,4
      DO 20 L=1,31
      IF ( X1(I,J,K,L) .EQ. 0.0 ) GO TO 20
      YY = YY + Y(I,J,K,L)*X1(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)*X1(I,J,K,L)
      PP = PP + WW
20    CONTINUE
      F(2) = YY - PP
      YY =0.
      PP = 0.
      DO 30 I=1,34
      DO 30 J=1,4
      DO 30 K=1,4
      DO 30 L=1,31
      IF ( X1(I,J,K,L) .EQ. 0.0 ) GO TO 30
      YY = YY + Y(I,J,K,L)*X2(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)*X2(I,J,K,L)
      PP = PP + WW
30    CONTINUE
      F(3) = YY - PP
      DO 60 I=1,34
      YY = 0.
      PP = 0.
      DO 50 J=1,4
```

```
      DO 50 K=1,4
      DO 50 L=1,31
      IF ( X1(I,J,K,L) .EQ. 0.0 ) GO TO 50
      YY = YY + Y(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)
      PP = PP + WW
50    CONTINUE
      F(I+3) = YY - PP - X(I+3)/SS1
60    CONTINUE
      DO 80 J=2,4
      YY = 0.
      PP = 0.
      DO 70 I=1,34
      DO 70 K=1,4
      DO 70 L=1,31
      IF ( X1(I,J,K,L) .EQ. 0.0 ) GO TO 70
      YY = YY + Y(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)
      PP = PP + WW
70    CONTINUE
      F(J+36) = YY - PP
80    CONTINUE
      DO 130 K=2,4
      YY = 0.
      PP = 0.
      DO 120 I=1,34
      DO 120 J=1,4
      DO 120 L=1,31
      IF (X1(I,J,K,L).EQ.0.0) GO TO 120
      YY = YY + Y(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)
      PP = PP + WW
120   CONTINUE
      F(K+39) = YY - PP
130   CONTINUE
      DO 150 I=1,34
      YY = 0.
      PP = 0.
      DO 140 J=1,4
      DO 140 K=1,4
      DO 140 L=1,31
      IF (X1(I,J,K,L).EQ.0.0) GO TO 140
      YY = YY + Y(I,J,K,L)*X3(I,J,K,L)
      WW = TWOW(X,X1,X2,X3,I,J,K,L,77)*X3(I,J,K,L)
      PP = PP + WW
140   CONTINUE
      F(I+43) = YY - PP - X(I+43)/SS2
150   CONTINUE
      PRINT*,'F(X)'
      PRINT*,F
      RETURN
      END
      FUNCTION TWOW (X,X1,X2,X3,I,J,K,L,N)
      DIMENSION X1(34,4,4,31),X2(34,4,4,31),X3(34,4,4,31)
```

```
      DIMENSION X(N)
      IF (J. EQ. 1.0) GO TO 95
      IF (K .EQ. 1) GO TO 100
85    W = X(1)+X(2)*X1(I,J,K,L)+X(3)*X2(I,J,K,L)+X(I+3)+X(J+36)+
     1X(K+39)+X(I+43)*X3(I,J,K,L)
      GO TO 110
95    IF (K .EQ. 1) GO TO 105
      W = X(1) + X(2)*X1(I,J,K,L) + X(3)*X2(I,J,K,L) + X(I+3) +
     1X(K+39) + X(I+43)*X3(I,J,K,L)
      GO TO 110
100   W = X(1)+X(2)*X1(I,J,K,L)+X(3)*X2(I,J,K,L)+X(I+3)+X(J+36)+
     1X(I+43)*X3(I,J,K,L)
      GO TO 110
105   W = X(1)+X(2)*X1(I,J,K,L)+X(3)*X2(I,J,K,L)+X(I+3)+X(I+43)*
     1X3(I,J,K,L)
C110   PRINT*,I,J,K,L,W
110   TWOW = EXP(W)/(1+EXP(W))
      RETURN
      END
```

```
      PROGRAM 2                                                  INVERT
(INPUT,OUTPUT,TAPE23,TAPE4,TAPE7,TAPE5=OUTPUT)
      COMMON /VALU/ X1,X2,X3
      COMMON /INV/ X,A,SS1,SS2
      DIMENSION A(77,77),AINV(77,77),WKAREA(3800)
      DIMENSION  X1(34,4,4,31), X2(34,4,4,31), X3(34,4,4,31)
      DIMENSION X(77)
      DIMENSION AIDEN(77,77)
      DATA N,IA,IDGT/77,77,1/
      DO 5 I=1,77
      DO 5 J=1,77
      AIDEN(I,J) = 0.
    5 CONTINUE
      CALL READ
      CALL FILL
      CALL LINV2F(A,N,IA,AINV,IDGT,WKAREA,IER)
C     PRINT*,IDGT
      DO 10 I=1,77
      PRINT*,A(I,I),AINV(I,I)
      WRITE(7,*) A(I,I),AINV(I,I)
   10 CONTINUE
C     PRINT*,A
      DO 20 I=1,77
      DO 20 J=1,77
      DO 20 K=1,77
      AIDEN(I,J) = AIDEN (I,J) + A(I,K)*AINV(K,J)
   20 CONTINUE
C     PRINT*,AIDEN
      STOP
      END
      SUBROUTINE READ
      COMMON /VALU/ X1,X2,X3
      COMMON /INV/ X,A,SS1,SS2
      DIMENSION X1(34,4,4,31),X2(34,4,4,31),X3(34,4,4,31)
      DIMENSION X(77), A(77,77)
      II = 1
      III = 1
   10 READ (23,15,END=20) A2,K,I,J,A3,L
   15 FORMAT (5X,F4.0,5X,I1,I2,5X,I1,27X,F5.0,16X,I2)
      IF (A3 .EQ. 99999) GO TO 10
      IF (III.NE.I) II=II+1
      X1(II,J,K,L) = A2
      X3(II,J,K,L) = A3
      X2(II,J,K,L) = X1(II,J,K,L) / X3(II,J,K,L)
      III = I
      GO TO 10
   20 READ (4,*) X, SS1,SS2
      RETURN
      END
      SUBROUTINE FILL
      COMMON /VALU/ X1,X2,X3
      COMMON /INV/ X,A,SS1,SS2
      DIMENSION X1(34,4,4,31), X2(34,4,4,31), X3(34,4,4,31)
      DIMENSION X(77), A(77,77)
```

```
      DO 5 I=1,77
      DO 5 J=1,77
      A(I,J) = 0.0
5     CONTINUE
      DO 10 I=1,34
      DO 10 J=1,4
      DO 10 K=1,4
      DO 10 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 10
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(1,1) = A(1,1) + WW*(1-WW)
10    CONTINUE
      DO 20 I=1,34
      DO 20 J=1,4
      DO 20 K=1,4
      DO 20 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 20
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(2,1) = A(2,1) + X1(I,J,K,L)*WW*(1-WW)
20    CONTINUE
      A(1,2) = A(2,1)
      DO 30 I=1,34
      DO 30 J=1,4
      DO 30 K=1,4
      DO 30 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 30
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(2,2) = A(2,2) + (X1(I,J,K,L)**2)*(WW*(1-WW))
30    CONTINUE
      DO 40 I=1,34
      DO 40 J=1,4
      DO 40 K=1,4
      DO 40 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 40
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(3,1) = A(3,1) + X2(I,J,K,L)*WW*(1-WW)
40    CONTINUE
      A(1,3) = A(3,1)
      DO 50 I=1,34
      DO 50 J=1,4
      DO 50 K=1,4
      DO 50 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 50
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(3,2) = A(3,2) + X1(I,J,K,L)*X2(I,J,K,L)*WW*(1-WW)
50    CONTINUE
      A(2,3) = A(3,2)
      DO 60 I=1,34
      DO 60 J=1,4
      DO 60 K=1,4
      DO 60 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 60
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(3,3) = A(3,3) + (X2(I,J,K,L)**2)*(WW*(1-WW))
```

```
60    CONTINUE
      DO 80 II=4,37
      DO 80 JJ=1,II
      DO 70 J=1,4
      DO 70 K=1,4
      DO 70 L=1,31
      I = II - 3
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 70
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      IF ( JJ .EQ. 2 ) THEN
      A(II,JJ) = A(II,JJ) + X1(I,J,K,L)*WW*(1-WW)
      ELSE IF ( JJ .EQ. 3 ) THEN
      A(II,JJ) = A(II,JJ) + X2(I,J,K,L)*WW*(1-WW)
      ELSE IF ((JJ.EQ.1).OR.(JJ.EQ.II)) THEN
      A(II,JJ) = A(II,JJ) + WW*(1-WW)
      ELSE
      A(II,JJ) = 0.
      ENDIF
70    CONTINUE
      IF ( JJ .EQ. II) A(II,JJ) = A(II,JJ) + 1/SS1
      A(JJ,II) = A(II,JJ)
80    CONTINUE
      DO 100 II=38,40
      DO 100 JJ=1,II
      J = II - 36
      IF ((JJ.GE.4).AND.(JJ.LE.37)) THEN
      I = JJ - 3
      DO 90 K=1,4
      DO 90 L=1,31
      IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 90
      WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
      A(II,JJ) = A(II,JJ) + WW*(1-WW)
90    CONTINUE
      A(JJ,II) = A(II,JJ)
      ELSE
      DO 95 I=1,34
          DO 95 K=1,4
          DO 95 L=1,31
          IF (X1(I,J,K,L) .EQ. 0.0 ) GO TO 95
          WW = TWOW (X,X1,X2,X3,I,J,K,L,77)
          IF ( JJ .EQ. 2 ) THEN
          A(II,JJ) = A(II,JJ) + X1(I,J,K,L)*WW*(1-WW)
          ELSE IF ( JJ. EQ. 3 ) THEN
          A(II,JJ) = A(II,JJ) + X2(I,J,K,L)*WW*(1-WW)
          ELSE IF ((JJ.EQ.1).OR.(JJ.EQ.II)) THEN
          A(II,JJ) = A(II,JJ) + WW*(1-WW)
          ELSE
          A(II,JJ) = 0.
          ENDIF
95        CONTINUE
      A(JJ,II) = A(II,JJ)
      ENDIF
100   CONTINUE
      DO 120 II=41,43
```

Credit Scoring Via Linear Logistic Models With Random Parameters

Kevin John Leonard

A Thesis

in

The Department

of

Decision Sciences and Management Information Systems

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

June 1988

ISBN 0-315-44881-4

ABSTRACT

Credit Scoring Via Linear Logistic Models
with Random Parameters

Kevin John Leonard, Ph.D.
Concordia University, 1988

Credit scoring models have received much attention in the literature. In this thesis, logistic regression models, some containing random effects, are employed to statistically represent or simulate the decision process of commercial loan officers. The use of random effects allows for an assessment of model adequacy, as well as for the quantification of variability among decision makers. Estimates of variance components are developed using the EM algorithm. These techniques are illustrated by an analysis of data from a major Canadian bank.

ACKNOWLEDGEMENTS

It would not be possible to undertake and complete a task as great as this without guidance and assistance from many people. I would like to acknowledge the help of a number of these people.

I would first like to thank Dr. T. Jerome Tomberlin, my supervising professor, for his insight, support, and leadership throughout the development of this dissertation. I could not have finished this thesis without his guidance.

I wish to thank Dr. Dale Doreen for introducing me to the topic of credit scoring, and especially for his encouraging support during the initial stages of the Ph.D. program.

I would also like to acknowledge the assistance of two other members of my supervising committee - Dr. Martin Kusy, with whom I first studied, and Dr. G. A. Whitmore. These two professors provided great help throughout the phases of the program, and accepted nothing less than the very best. I thank them.

I would like to thank Dr. D.G. Kinnear and Dr. D.G. Hampson for always being there when I needed them most. For this I'am eternally grateful.

I would like to thank the management and the employees of the bank which provided the data; without their cooperation, this thesis could never have been written.

I thank the Computer Center, and especially Ms. Tan Bui, who worked on the programming with me and gave countless hours. I would also like to extend my gratitude to Mr. Pierre Balthazard who provided computer assistance.

To the department of Decision Sciences and Management Information Systems (the QM department) for being very supportive and to Ms. Toni Anselmi for being so helpful with all those "tiresome" figures and tables, I extend my appreciation.

To the members of my family who supported me throughout and who, I know, celebrate the completion of this degree as much as I, I say to them - "it was worth it." I especially mention my parents, John and Rita, who taught me to be the best that I could be. I thank them for always being there.

I wish to thank Teresa Heffernan who spent many hours proof-reading the final document; without her help this dissertation would not have been completed.

And, finally, to my best friend, my companion, my wife, Sandra, I cannot express what having her there has meant to me. This paper emanates from the love she has given me.

This thesis is dedicated to the memory and the spirit of two great Canadians whose odysseys have marked the beginning and the end of my work in the Ph.D. program :

TERRY FOX

RICK HANSEN

<u>Table of Contents</u>

Page

Page

LIST OF TABLES AND FIGURES

CHAPTER ONE

INTRODUCTION

The process of modelling the variables important in the extension of credit is referred to as "credit scoring." This modelling process which often uses statistical methodology (such as discriminant analysis) has been carried out by banks and financial institutions. Unfortunately, the results from these studies are often withheld for reasons of confidentiality. Credit scoring research has also been undertaken by academics; however, two major shortcomings have been the unavailability of quality secondary data and the inability to obtain new data. Consequently, there is a need for more indepth research in this area.

In this thesis, we study actual bank data using a number of statistical techniques in order to address these credit scoring problems. The major areas of research are described in detail below. Emphasis is placed on the contributions to the literature in each of these major areas.

1.1 Dissertation Topic

The three principal areas of investigation are the following:

(1) Credit Scoring Problem

The credit scoring problem has been widely investigated. (See Section Two for a full discussion of credit scoring.) Here, the outcome of small business loan applications, whether approved or declined, is investigated. More specifically, an analysis of actual credit files from one of Canada's largest credit institutions is presented. The creation and examination of credit scoring models based on analysis of data from actual bank files constitutes the major empirical contribution.

(2) Logit Regression Models with Random Effects

The topic of logit regression models has also received a great deal of attention in the literature (see Chapter Four). There are statistical packages presently available which perform logit regression and yield maximum likelihood estimates for the parameters (for example, BMDP-83). However, these packages are only available for fixed effects models. Wong and Mason (1985) have applied logit models where the slope and effects parameters are considered random. In this thesis, similar models with a mixture of random and fixed effects (mixed models) are considered and applied in a credit scoring setting. This application requires significant mathematical formulation and software development. These models provide empirical Bayes estimates of the probabilities for the success of credit applications. This development constitutes the majority of the theoretical contribution of this dissertation.

(3) Comparisons with Other Statistical Techniques

The final area of investigation involves the analysis of the performance of three statistical techniques in modelling the credit scoring problem, namely, linear discriminant analysis, logit regression and logit regression with random effects. (Both linear discriminant analysis and fixed effects logit regression programs are readily available in statistical packages, and, as such, do not require software development.) The comparisons involve goodness-of-fit testing, the evaluation of predictive ability using cross validation techniques (e.g., holdout samples, bootstrapping), the examination of the statistical significance of the models, the analysis of statistically significant independent variables, and the analysis of residuals. As well, the predictive ability of these three statistical techniques is evaluated in comparison to naive models where all cases are simply classified into the modal group or groups. Analysis of these techniques results in significant contributions to both theoretical and empirical research.

1.2 Dissertation Objectives

The primary objective of this dissertation is to introduce a statistical methodology for the purpose of analyzing the credit scoring problem. Since this particular methodology is new to credit scoring, many restrictions are inherent which limit applicability to other studies. (These types of restrictions are discussed in detail throughout this thesis.) The important point for consideration is the application of new, state of the art techniques to an old problem. It is hoped that, as a result of this paper, future research will examine and extend the models presented here to create more realistic and reliable credit scoring models.

We develop a decision making model which will aid in the prediction of the outcome, approval or decline, of small business loan applications for one division of a major Canadian bank. A clarification of the term "decision making model" is warranted. A model used for decision making can be described as either normative or descriptive (Dawes and Corrigan, 1974). A normative model aids the decision maker in reaching a good decision. A descriptive model, on the other hand, represents the decision maker's behaviour. The model presented in this thesis is normative in nature and is used as part of a larger decision making process. An examination of a descriptive model, a model of the bank manager's behaviour, will be reserved for future research.

Decision models are also often discussed in the context of decision support systems (DSS). Although the models presented here are not created in an attempt to build a decision support system for banks, they could be employed within such a system. These models can be better classified as "operational level" models within the broader framework of the DSS model. For an elaborate discussion of DSS for banks and the place of credit scoring models therein, see Sprague and Watson (1976).

In summary, we present normative credit scoring models which could assist commercial loan officers in deciding whether or not to extend credit.

CHAPTER TWO

THE CREDIT SCORING PROBLEM

In this chapter, a review of the literature pertaining to credit scoring models is presented. Some of the landmark developments are described, but emphasis is placed on the direction and objectives of present literature.

2.1 Conceptual Underpinnings of Credit Scoring

In today's marketplace, the term "credit" is used in a number of contexts. It is used to describe purchases, types of cards, and even financial institutions (credit unions). In this paper, credit is defined as an amount of money (line of credit) which a customer has access to for a specific time period.

In more detail, we consider the case of banks lending money to their clients. When a customer borrows this money, he is trading off the future cost of this money for the benefit of having the money immediately.

Banks are not the only businesses which extend credit to their customers. In almost all industries, companies sell to their customers on credit. The balances outstanding on their accounts are paid at a future date, often at the end of the month. When companies extend this credit they are incurring a risk. There is the possibility that the purchasing company will not repay its debt obligation. This could become very costly for the seller who may incur legal fees, suffer opportunity costs, and, possibly, be forced to write-off the total outstanding debt. Thus, it is in the best interest of all credit granting firms to examine the creditworthiness of their customers before offering them a line of credit. (A detailed study analyzing the credit granting procedure for business firms is presented in Bierman and Hausman [1970].)

Now the banking industry is examined more closely. Commercial loan applicants require money to satisfy present financial needs. Banks must evaluate the ability of these firms to repay their financial obligations according to the agreement established between the respective parties. This evaluation process can be carried out using credit scoring models. (Credit evaluation can be done in a number of ways, however, in this study, we concentrate solely on credit scoring models.) Based on statistical analyses of historical data, certain financial variables are determined to be important in the evaluation process of the customer's financial stability and strength. Hence, information on these variables is then obtained for new bank customers. This information is summarized and a point system is used where the different variables have different weights. An overall score is produced by adding these weighted scores. If this overall score is above a predetermined cut-off point, the loan applicant receives a certain line of credit. If not, the applicant is denied credit.

The reasons for the creation of a financial credit scoring model can be summarized as follows:

(i) to quantify the mechanical procedures involved in credit scoring; and

(ii) to identify the variables which are important in the credit evaluation process, where this identification provides feedback to the credit managers.

This is a very brief introduction to financial obligations, credit and credit scoring. For a more detailed discussion, see, for example, Reichert, Cho and Wagner (1983).

The different characteristics of credit scoring models present in the literature are described below. There is no general consensus concerning the choice of predictor variables, how these variables should be represented in the model, nor which statistical techniques should be implemented for data analysis (Altman, 1968; Orgler, 1970; Reichert, Cho and Wagner, 1983). Six areas of contention are examined. In order to better illustrate the discussion, a tree diagram is provided.

2.2 Loan Evaluation vs. Bankruptcy Prediction

Much of the research in credit scoring is focussed on two areas:

(i) predicting bankruptcy, and

(ii) predicting the outcome of a loan.

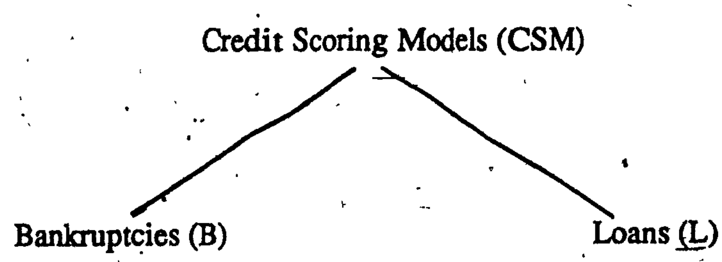


Figure 2.2.1

One of the first significant investigations into the field of credit scoring was performed by Altman (1968), in which the prediction of corporate bankruptcies was described. He incorporated a scoring model using linear discriminant analysis to distinguish between bankrupt and non-bankrupt companies. The independent variables are all financial ratios and are discussed later. Altman defined bankrupt firms as "firms [which] are legally bankrupt and either [have been] placed in receivership or have been granted the right to reorganize under the provisions of the National Bankruptcy Act".

Other prominent work concerning credit scoring for the prediction of business failures has been conducted by Deakin (1972), Edmister (1972), Blum (1974), and further work by Altman (1973,1977).

Credit scoring has also been applied in the evaluation of loans. More specifically, models for the evaluation of loans can be segmented into two groups:

(i) models for the prediction of the Outcome of the Loan Repayment - OLR (the loan is repaid in full or there is a default); and

(ii) models describing the actual decision process of the loan officer with respect to the Outcome of the Loan Application - OLA (the loan is approved [an approval] or the loan is declined [the banks refer to this as a "decline".]).

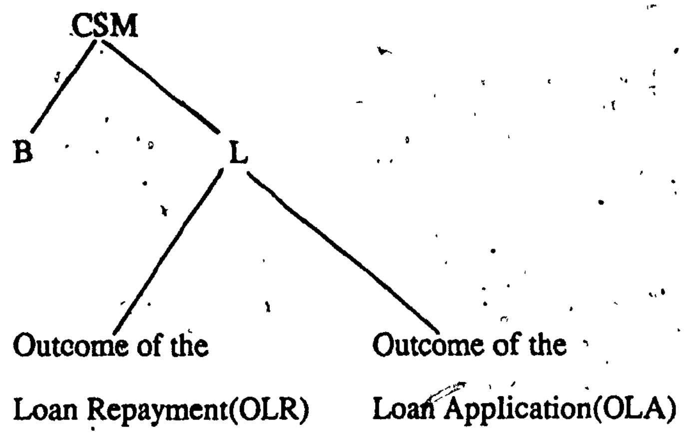


Figure 2.2.2

One of the first major investigations into the prediction of the loan repayment (OLR) was conducted by Orgler (1970). He developed a multiple regression model to discriminate between good and bad commercial loans. The dependent variable was a dichotomous variable defined as follows: "Any loan which was criticized by a bank examiner was considered as bad while any loan which was evaluated but not criticized was classified as good." Classification was not based on the actual pay-off (or charge-off) of the loan because many of the loans which were eventually paid off were either not profitable or incurred a loss due to delays and penalties.

Further research relating to the prediction of the outcome of loan repayments was conducted by Orgler (1971), Edmister (1971), and Wiginton (1980).

The second group of models represented in Figure 2.2.2 consists of those used in predicting the outcome of the loan application (OLA). In general, loan officers decide whether loan applicants will receive credit or not. Most applicants, as well as the executives of the credit institutions, are interested in determining which criteria variables are prominent in the loan officer's decision. If loan applicants had knowledge of these important variables, then their applications could be submitted in a manner which concentrates on the appropriate criteria variables. This would presumably result in a higher probability of success, i.e., the loan would be granted).

Cohen, Gilmore and Singer (1966) performed an elaborate study which analyzed the decision process followed by the loan officer (OLA). The objective of their research was to describe in a "careful, analytical fashion the way in which banks currently decide whether to grant a proposed business loan." The researchers attempted to model present behaviour; they did not try to provide an "optimal" model. The authors separated the loan officer's decision criteria into eight main steps:

(i) evaluate the status of the firm's customer relationship;
(ii) evaluate new customer relationship;
(iii) perform credit evaluation;
(iv) examine legal and policy restrictions;
(v) appraise the loan's purpose, amount, maturity, pay-back and security;
(vi) review the detailed recommendations;
(vii) record analysis and recommendations; and
(viii) follow-up and review.

The information from all of these steps was combined using a model which incorporated simulated data in order to arrive at the decision of the loan officer.

Subsequent research on modelling the decision process for the loan application (OLA) was conducted by Doreen and Farhoomand (1983), and Reichert, Cho and Wagner (1983).

2.3 Commercial Loans vs. Consumer Loans

The study of credit scoring for loans (both OLR and OLA) can be further segmented into two categories:

(i) consumer loans, and
(ii) commercial loans.

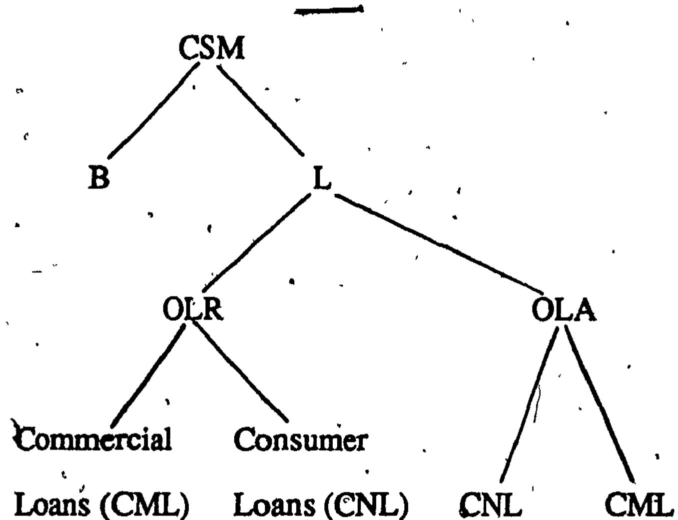


Figure 2.3.1

Credit scoring models were initially employed in consumer credit studies (Durand, 1941). Consumer credit evaluation has received a great deal of attention both in academia and in business practice. In fact, an overwhelming majority of banks and credit card companies presently use some type of checklist developed through a "credit" analysis model in order to evaluate consumer credit (Wiginton, 1980).

One of the more important studies of consumer loans was performed by Orgler (1971), in which he concentrated on modelling the OLR. Variables pertaining to three major areas were incorporated:

(i) personal information on the borrower (such as monthly income, monthly rent and number of years employed by current employer);

(ii) subjective information provided by the loan officer (such as credit rating); and

(iii) performance data (e.g., number of times the customer was overdue in past dealings with the institution).

Subsequent research on the topic of consumer loans was completed by Wiginton (1980) and by Reichert, Cho and Wagner (1983).

The second area, commercial loans, has also received much attention in the literature. However, these credit scoring models have not received the same widespread implementation as have their consumer counterparts. One explanation for this lack of implementation is the absence of agreement among researchers regarding which independent variables to include in the model. The following quote by Orgler (1970, p.436) describes this situation:

> For several reasons, it is difficult to apply the methodology used in credit scoring consumer loans to commercial loans. First, commercial borrowers do not belong to large homogeneous populations as do customers for consumer credit. This lack of standardization presents a problem in obtaining sufficient data for a statistically significant study. Second, there are substantial variations among commercial loans with respect to their size, terms, collateral types and payment procedure, all of which are relatively uniform in the case of consumer loans. Finally, there is a lack of reliable up-to-date financial data on small commercial borrowers, and particularly on those who defaulted on their loans.
>
> Because of these difficulties, it is impossible to develop a general scoring model for commercial loans. Instead, it is necessary to develop individual models for small commercial borrowers in each industry, provided that sufficient data are available. This limitation may be one of the reasons for the lack of analytical models in this area.

See Section 2.8 for further discussion of Orgler's comments.

2.4 Small Business Loans vs. Commercial Loans

As described by Orgler (1970), the field of commercial loans consists of a diverse collection of companies ranging from small businesses to large international corporations. The task of creating one general credit scoring model is complicated by this diversity. For example, the predictor variables may vary in accordance with the size of the firm. In large firms, a significant independent variable could be the percentage change in share price over some time interval. On the other hand, in small businesses, which are commonly privately owned, this independent variable would not be available. Hence, predictor variables are peculiar to the size of the firm. One solution is to limit the creation of credit scoring models to small homogeneous groups of firms with common predictor variables. One such homogeneous group is that of small businesses.

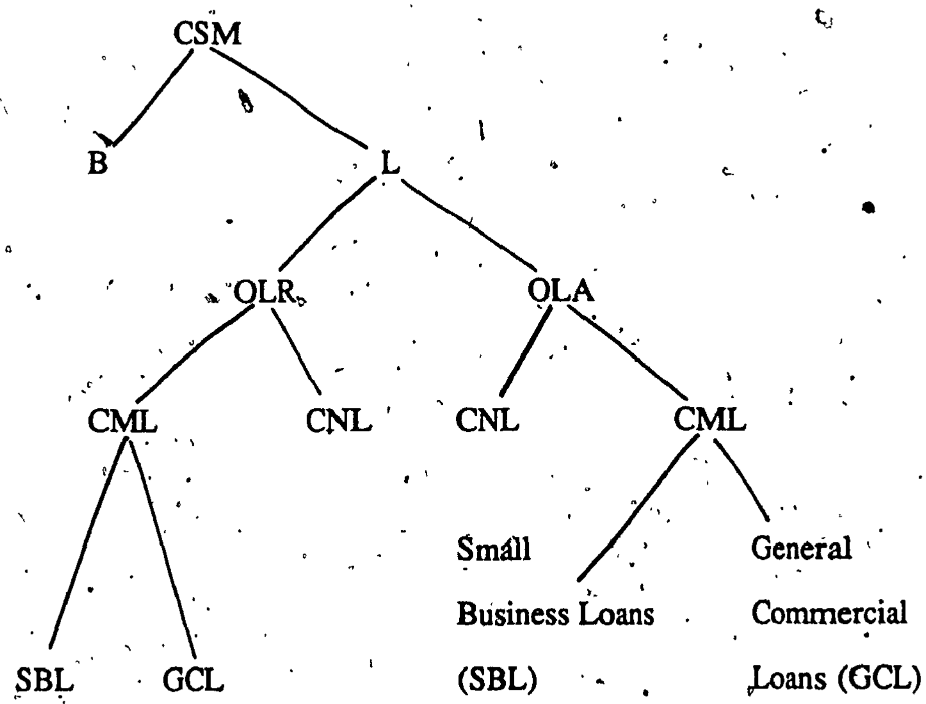


Figure 2.4.1

There is no universally accepted definition of a small business. It is generally understood that small businesses are companies which are usually operated by an owner/manager and which are not dominant in their field of operation. Generally, sales figures are not given because the dollar values of today can become obsolete over time due to inflation. Consequently, the designation of "small business" includes a variety of valid definitions involving sales amounts and other monetary value limits (Szonyi and Steinhoff, 1979).

Edmister (1971,1972) introduced credit evaluation for both the failures in small business and for the outcome of small business loan repayments. In the case of modelling the OLR, a linear discriminant model was derived using financial ratios as the independent variables. However, little emphasis was placed on the unique situation of small businesses.

In 1983, Doreen and Farhoomand created a credit decision model for small business, and predicted the OLA using linear discriminant analysis. The authors considered independent variables peculiar to small businesses (such as owner/manager initiative). They explained the importance of examining small business loans as a special group of commercial loans (1983, p.19):

> First, the evaluation of a small business loan requires a great deal of judgement by the credit manager of the lending institution. In contrast, the commercial loans for larger companies can be more readily evaluated by using credit histories and the financial ratios of the firm.
>
> Second, although small business loans are usually small in magnitude, a credit manager has to spend as much time processing them as he does a commercial loan. As a result, the opportunity cost of evaluating a small business loan is much higher than the cost of evaluating a commercial loan.
>
> Third, because of the peculiarity of small business, the personal characteristics and the managerial capabilities of the owner/manager play a more important role in the loan evaluation process.
>
> Because of these differences, we do well to differentiate between commercial loans and small business loans.

2.5 Primary Data vs. Secondary Data

The term "primary data" refers to data that are directly obtained for the purposes of the study involved. For example, researchers prepare a checklist of relevant variables and, often through the cooperation of a credit institution, collect new data for analysis. The term "secondary data" refers to existing data available in the institution's files. Much of the early work in credit scoring (for either bankruptcies or loans) is based on secondary data (for example Altman, 1968; Orgler, 1970).

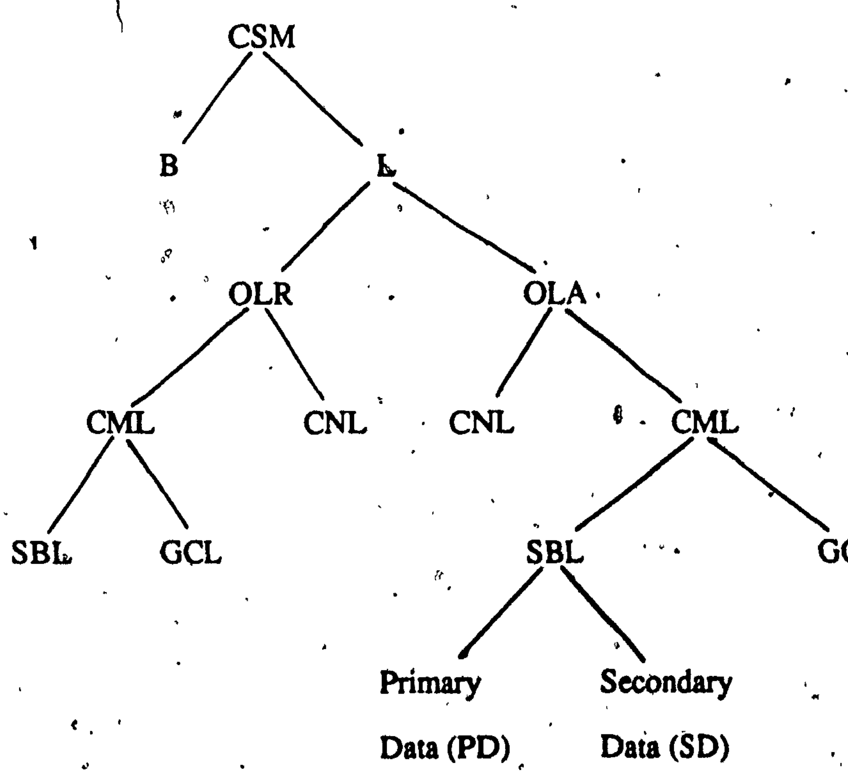


Figure 2.5.1

A major advantage in using primary data is that one can collect data on the variables desired. Specially designed questionnaires can be used in the compilation of information on variables which the researchers deem important. A major disadvantage of this method is the poor response rate when collecting new data. This is greatly amplified when the research is undertaken by parties outside the financial institution involved. There can be resistance to compliance and a large reduction in the response rate. In the Doreen and Farhoomand study (1983), for example, the response rate on collecting primary data was only 6.33% (38 out of 600 questionnaires were returned). However, if the objective of a credit scoring model is to accurately model or simulate the methods which are presently in use rather than change current evaluation methods, then sufficient information may exist in the files of the credit institutions. A model of the decision process can be made based on secondary data.

An underlying assumption is that all pertinent information relating to the decision process is contained in the files. Unfortunately, in this study, because interviews with credit managers were not permitted, as is discussed in Chapter Three, there is no way to verify this assumption. Thus, it is possible that the information in the files is incomplete.

Although the data may not be complete, the data which are available may focus on the variables which were used in the decision process. In other words, the manager's records may concentrate either on the negative points, in the case of a decline, or on the positive points, in the case of an approval, as opposed to on all relevant information. This focus may tend to amplify the differences between the approvals and the declines. This would then tend to make the cases which are marginally rejected or approved appear to be clear cut decisions from the start. This amplification of the differences should be considered when the analysis is performed.

2.6 Subjective vs. Objective Predictor Variables

Most of the predictor variables used in credit scoring models developed to date have been objective. Here, objective data are referred to as factual data, which are obtained from official records and documents.

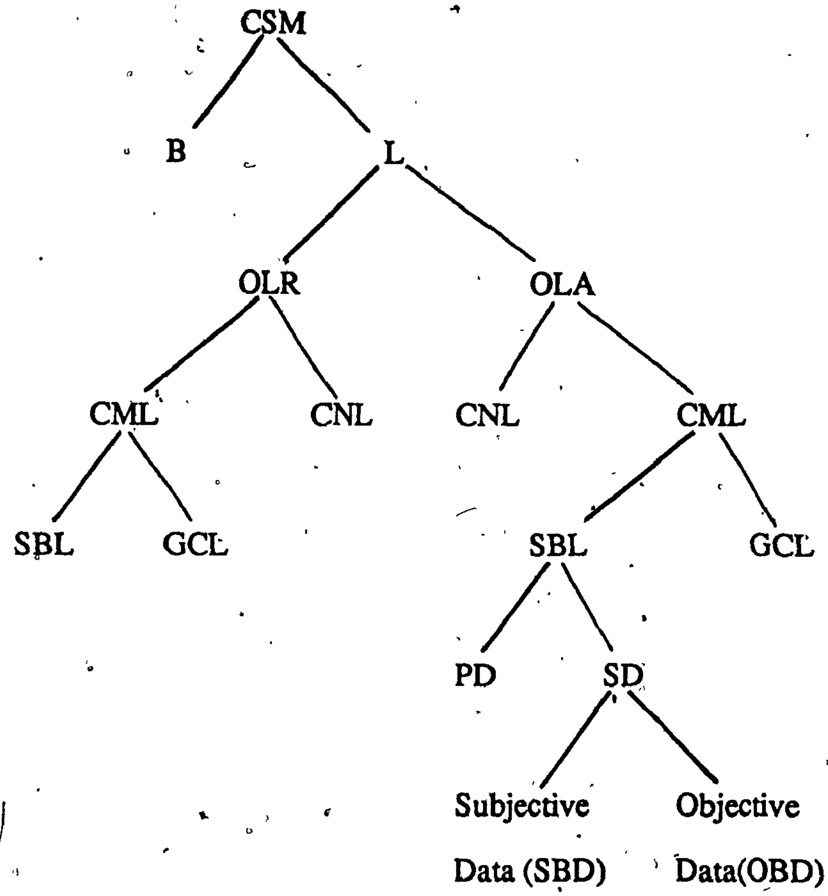


Figure 2.6.1

Altman (1968) used financial ratios in his discriminant model for the prediction of the OLR for commercial loans (in general). He concluded that the following five ratios were good predictors:

(i) working capital / total assets;

(ii) retained earnings / total assets;

(iii) earnings before interest and taxes / total assets;

(iv) market value equity / book value of total debt; and

(v) sales / total assets.

In the case of predicting the OLA for consumer loans, Reichert, Cho and Wagner (1983) also used objective variables in their discriminant model. The six predictor variables were:

(i) credit rating;

(ii) relative debt load;

(iii) job age;

(iv) residence age;

(v) indirect-direct lending; and

(vi) number of payments in contract.

In their study on modelling the OLA for small business loans, Doreen and Farhoomand (1983) examined subjective data (in addition to objective data). Subjective data are defined as data that are measured judgmentally by the credit manager when collected and recorded. Some of the variables included were owner/manager experience, owner/manager communication skills, and competitive situation in the industry. Subjective variables are usually much harder to incorporate into the model than objective variables due to this judgmental factor. This leads to a lack of standardization in the data. In this thesis, the data consist solely of objective information obtained from bank files.

It is possible that inclusion of subjective variables would lead to a better predictive model, however, this is left for future research.

2.7 Discriminant Analysis vs. Other Statistical Techniques

In the area of credit scoring, multivariate normal linear discriminant analysis has been the statistical method favoured by most authors (Altman, 1968; Doreen and Farhoomand, 1983; Reichert, Cho and Wagner, 1983). As Wiginton (1980) states, "most models are based on the concept of scoring by use of weights usually determined as statistically significant coefficients of some linear statistical model, frequently the linear discriminant model." Unfortunately, due to the nature of credit scoring data, the assumptions required for applying discriminant analysis are often not realistic and other statistical methods should be considered (Eisenbeis, 1977). (Refer to Chapter Four for a detailed discussion of this area.)

An alternative statistical method is multiple regression (Orgler, 1971). The dependent variable for studying the OLR is a zero-one variable, which indicates good or bad loans. In the statistical literature, this is often referred to as "binary regression" (Green, 1978). However, this technique is equivalent to discriminant analysis when only two groups are involved, and it provides the same results (Orgler, 1971). Consequently, the difficulties in using discriminant analysis also extend to multiple regression.

Another alternative statistical method for credit scoring is logit regression. (This technique is discussed further in Chapter Four.) This method has received a great deal of attention in recent years and has proven to be effective in a variety of settings.

2.8 Direction for Research

In this chapter, many characteristics of credit scoring models have been introduced and discussed. There are many variations in the models proposed by the researchers, particularly in relation to one or more of the six broad areas discussed above.

For the purposes of this dissertation, only one subset of the credit scoring literature is examined. More specifically, we study the decisions of loan officers in relation to small business commercial loans at one Canadian bank (modelling the OLA). The research concentrates on objective data presently amassed by credit managers for their credit evaluation (secondary data). This particular subset of the total credit scoring scenario is highlighted in the final tree diagram (the bold path), Figure 2.8.1.

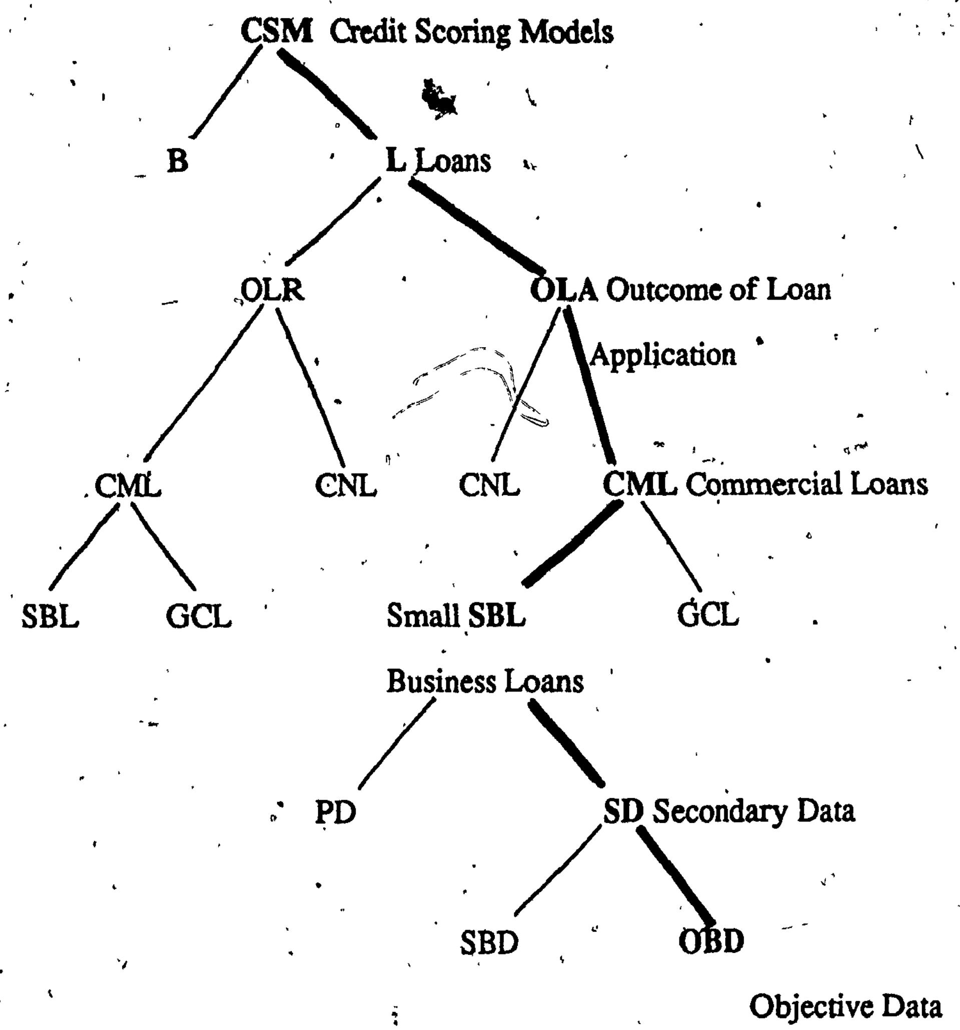


Figure 2.8.1

At this point, we refer back to Orgler's comments (Section 2.3) on the viability of creating credit scoring models for commercial loans. Following his suggestions, we concentrate on building a credit scoring model for a specific group of "small commercial borrowers." We have not attempted to build a universal commercial loan model but rather a model which applies to the subset of commercial loan applicants described above.

In conclusion, it should be noted once again that the purpose of the credit scoring model presented in this paper is to complement the decision making process that is already in use - not to replace it. In this view, the creation of a credit scoring model is a worthwhile endeavour. The model can perform routine calculations which now occupy valuable time of the credit manager; it can easily "flag" missing information when the manager consults the model; and it can be used as a "second opinion" - often supporting the original decision of the credit manager (validation). If the two opinions differ, this will, at the very least, suggest that further investigation into that particular loan application is warranted.

In short, the credit scoring model will: (i) mechanize and standardize some of the evaluation process; and (ii) provide feedback to the credit managers.

CHAPTER THREE

DATA COLLECTION

In this chapter, we describe the collection process used to organize and complete the data set. This chapter concludes with a brief discussion on the model formulation to be employed in this thesis.

3.1 Research Methodology

The credit evaluation process currently in practice at one regional division of a major Canadian bank is investigated. As the first step in this investigation, the actual bank files pertaining to small business loans for the period from January 1, 1985, to December 31, 1985 were examined and data therefrom collected. Due to the restrictions of confidentiality, the investigation was limited to new customers at one division over a period of one year.

Secondly, the decision making process at the regional division office was studied. It was determined that credit decisions were made, in effect, by two separate levels of management:

(i) the Division office - Division Managers, and
(ii) the Branch offices - Senior Branch Manager.

Every level of management (both at the division office and at the branches) of this particular Canadian bank has decision making power dictated by the dollar value of the loan application. Any dollar value below the manager's "decision interval" would be handled by his subordinates. A loan application for an amount above this "interval" would be forwarded to the manager's supervisor for approval or rejection.

In this study, all applications examined were for a dollar amount above the Senior Branch Management upper limit (which varies from branch to branch) and below the Senior Division Manager's upper limit. In this instance, the protocol is for the Senior Branch Manager to screen the applications and forward only the ones he "recommends" to the division office. Thus, the final decision concerning <u>all</u> the cases recommended is made by an individual at the division office. All files below this interval are kept at the branches; permission to review these files was not granted. Those above the Senior Branch Manager's upper limit are sent to the Corporation Headquarters.

It should be noted that the upper limit of the amount of the loan application was restricted to one million dollars. (All dollar values are in Canadian currency.) For reasons explained in Chapter Two, we are primarily interested in small business loans. However, this bank does not have any official "small business" designation. Thus an arbitrary cut-off point of $1 million was adopted. To further insure that the analysis was restricted to small business loan applications, only firms with gross sales figures under ten million dollars were retained in the sample.

The diagram below illustrates the actual decision process:

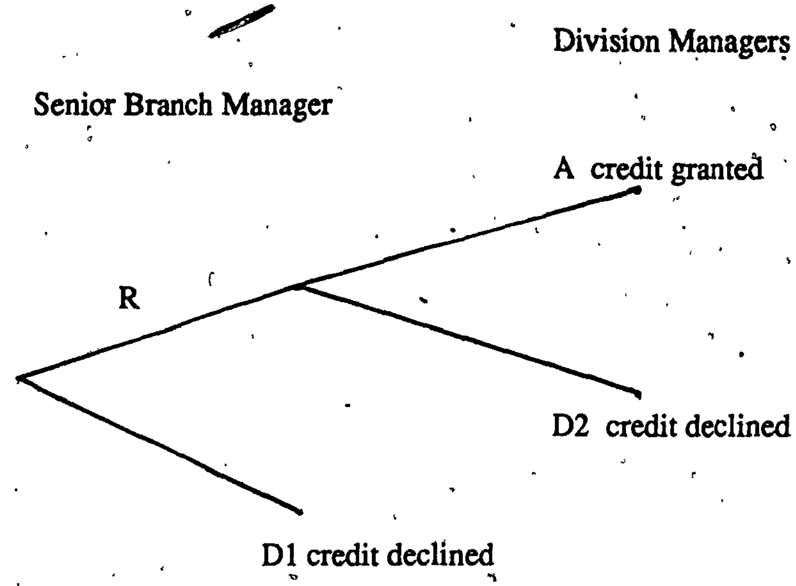


Figure 3.1.1

where R represents a loan application recommended by the Senior Branch Manager for approval, D1 represents a loan application declined at the Senior Branch Manager level, A represents a loan application approved by Division Managers, and D2 represents a loan application declined by Division Managers.

At first, a loan application is filed at any commercial branch (Commercial Banking Center - CBC) of this bank. The Senior Branch Manager then peruses the application and if the dollar amount is within his credit interval, he decides whether it is approved or declined. If it is above his credit limit, it is forwarded to the division office with his recommendation (R) for approval. The

Division Managers will then decide whether the application is approved (A) or declined (D2). It is possible for a Senior Branch Manager to make a decision concerning a loan above his credit limit. (This limit can vary from branch to branch depending on the capabilities of the specific manager involved.) If the Senior Branch Manager judges that the loan applicant does not have the necessary credentials, then he will not forward the application regardless of the dollar amount, and the loan application will be declined (D1). Thus, the Senior Branch Managers are involved in all loan applications filed at the regional division office. It should be noted that the analysis described here pertains to those applications for amounts above the Branch Managers' credit limits.

3.2 Outline of the Data Collection

In the Cohen, Gilmore and Singer (1966) study (Section 2.2), the researchers discussed eight main components which influence the decision of the loan officer (OLA). Due to three restrictions outlined below, analysis of the research presented here is limited to three of these eight components:

(i) evaluate new customer relationship;
(ii) perform credit evaluation; and
(iii) appraise the loan's purpose, amount, maturity, pay-back, and security.

The first of these restrictions is that this study is limited to customers who are new to the bank. This eliminates the need for any evaluation of the firm's history with the bank (and any subsequent follow up and review).

Second, interviews with the Senior Branch Managers (or any other credit manager) were not permitted. Therefore, the data collected relate solely to objective information available in the bank files. Subjective information was often not committed to paper, and when it was this information was not standardized.

This second restriction leads directly to a third one. To reiterate, only data from the bank files were obtained. The bank files contain information on possible clients who have enough initial credibility for the credit manager to commit the application to paper. It is possible that an informal screening process of loan applicants occurs even before any data are officially recorded and a file is opened. Information concerning applicants who are "informally" rejected could not be obtained. Thus, inferences from this study will be limited to those loan applicants who have enough credibility to obtain a first interview and to have a bank file created.

3.3 Investigation of Bank Files

Bank files at a single division office were examined. Information was recorded with respect to 278 loan applications that were either approved (A) or declined (D2) by the Division Managers. There were 252 loan applications approved and 26 declined. These data pertain solely to the decision process followed by the Division Managers (whose identities were not recorded). It does not contain any direct information about the decision process followed by the Senior Branch Managers. However, these applications were previously screened and subsequently recommended for approval (R) by the Senior Branch

Manager involved. Thus, in order to validate conclusions drawn, information was obtained concerning the decision process effective at the Senior Branch Manager level.

Consequently, bank files from four branches (selected by the bank) were forwarded to division offices and a total of 61 more applications was recorded. All of these applications represented credit declines (D1) at the Senior Branch Manager level. (The identity of the Senior Branch Manager involved in the final decision was recorded.) As a result, data were obtained pertaining to the decision process at both levels - division and branch. The following diagram illustrates the origin of the data.

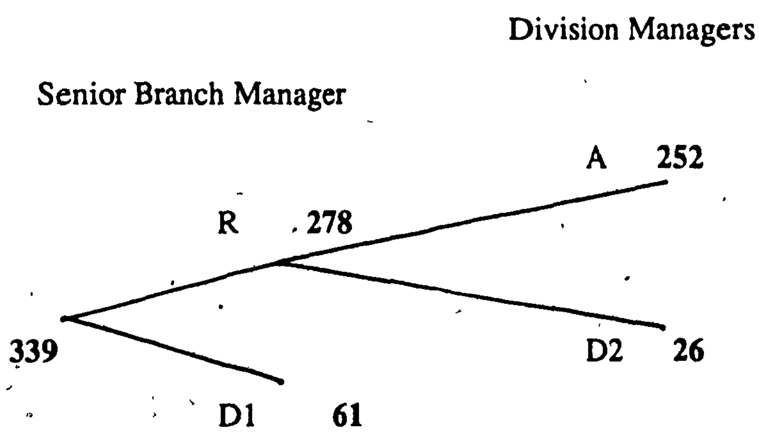


Figure 3.3.1

A total of 339 bank files was examined.

(Remark: It was this author's intent to have both the samples - at the division and at the branch level - contain all loan applications at these levels at this bank over the one year period. However, there were no controls in place to insure that the files for all loan applications were forwarded and subsequently analyzed. Division Management requested that the selected branches forward <u>all</u> loan application files to the division office, but there was no investigatation regarding their compliance to this request. All conclusions and inferences are subject to the accuracy of this assumption.)

The data consisted of the following variables:

(i) company name (NAME);
(ii) company number (ACC);
(iii) outcome of application - approved or declined (OUT);
(iv) amount of the loan application (AMT);
(v) total credit position of the company (CRE);
(vi) new vs. existing company (NEW);
(vii) branch of bank (where the application was filed) (BRN);
(viii) type of loan (operating/term/small busine loan) (TY);
(ix) financial statement type (past records/pro-forma/none) (ST);
(x) sales (SA);
(xi) gross profit (GP);
(xii) operating profit (OP);
(xiii) bad debts (BD);
(xiv) depreciation (DEP);
(xv) income surplus (loss) transferred to equity (INCO);
(xvi) total net worth (TNW);
(xvii) net working capital (WC);
(xviii) private/other company (TF);
(xix) number of owners of firm (NO); and
(xx) owners guaranteed security in the company (SEC).

In order to give an insight into the nature of the data, we divide the variables into three types - identifiers, numerical, and categorical.

Identifiers	Numerical	Categorical
NAME	AMT, OP, TNW	OUT(2) *
ACC	CRE, BD, WC	NEW(4)
	SA, DEP, GP	BRN(48)
	INCO, SEC	TY(4), ST(3)
		TF(2), NO(2)

* The numbers in the parentheses represent the number of categories.

In short, we are modelling the outcome of the decision process of the loan officer, evident at one division office, from Senior Branch Manager responsibility to Division Management. The dollar amounts of the loan applications range from the Senior Branch Manager's upper limits to a ceiling of $1,000,000.

3.4 Description of the Raw Data

In all, 339 cases from 48 branches were examined. The 61 declines at the branch level involved only 4 branches. The remaining 278 cases involved 252 approvals and 26 declines. A complete breakdown of the observations by each branch is given below.

Table 3.4.1

Distribution of Observations by Branch

Branch	Declines (D2)	Declines (D1)	Approval (A)	Total
1	3	0	22	25
2	0	0	1	1
3	0	0	1	1
4	0	17	11	28
5	0	0	2	2

6	1	0	9	10
7	0	0	1	1
8	0	0	1	1
9	0	0	1	1
10	0	0	4	4
11	7	34	24	65
12	0	0	3	3
13	0	0	1	1
14	0	0	2	2
15	0	0	6	6
16	0	0	1	1
17	0	0	4	4
18	0	0	7	7
19	0	9	4	13
20	0	0	1	1
21	1	0	14	15
22	0	0	1	1
23	0	0	1	1
24	0	0	11	11
25	0	0	1	1
26	0	0	4	4
27	2	0	16	18
28	0	0	1	1
29	0	0	1	1
30	1	0	5	6
31	0	0	11	11
32	0	0	2	2
33	1	0	3	4
34	1	0	1	2
35	0	0	4	4
36	1	0	7	8
37	1	0	5	6
38	1	0	0	1
39	0	0	1	1
40	1	0	16	17
41	0	0	5	5
42	1	0	4	5
43	0	0	6	6
44	0	0	3	3
45	0	0	2	2
46	1	0	6	7
47	2	0	2	4
48	0	0	3	3
Total	26	61	252	339

The chart illustrates that 106 out of 339 cases (or 31.27%) were generated by only three branches (4,11,19). There were many branches (31) which had either approvals or rejections, but not both. As well, many branches had a limited number of observations. This occurrence plays a large role in the data analysis and is discussed at length in Chapter Five.

The range of values for the numerical variables is given below.

Numerical Variables	Mean	Standard Deviation	Range ($ 000)
AMT	203.97	229.05	5 to 1000
OP	81.17	187.77	-645 to 1674
TNW	174.95	421.14	-1700 to 5100
CRE	310.39	231.46	7 to 10000
BD	5.56	13.16	0 to 62
WC	118.96	442.64	-202 to 2392
SA	1105.171	394.05	0 to 10000
DEP	14.49	21.53	0 to 410
GP	440.96	484.02	-158 to 9900
INCO	61.44	191.58	-582 to 1025
SEC	279.14	840.56	0 to 5800

Finally, a breakdown of the codes for the categorical variables is provided. The description of each category, as well as the number of observations therein, is given.

Categorical Variable	Coding (Number of Observations)	Explanations
OUT	0 (87)	- loan declines
	1 (252)	- loan approvals
NEW	1 (101)	- new firm with bank history
	2 (159)	- new firm with no history
	3 (27)	- existing firm with history
	4 (52)	- existing firm with no bank history
TY	1 (138)	- operating loan
	2 (54)	- loan for equipment
	3 (51)	- loan for land/building
	4 (96)	- a combination of all 3
ST	1 (108)	- proforma statements
	2 (221)	- past statements
	3 (10)	- not available
TF	1 (305)	- private firm
	2 (34)	- other type of firm
NO	1 (145)	- single owner
	2 (194)	- more than one owner.

Missing observations (see Section 5.6) only posed a problem for the total net worth variable (TNW); in fifty-six (out of 339) cases this value was absent.

See Section 5.2 for a further discussion of the format of the data.

3.5 <u>Model Specification</u>

At this stage of the presentation of the data, it is appropriate to introduce the formulation of the model that is presented in Chapter Five. Due to the peculiarities of the data set described above, it is important to justify here the modelling approach that is used.

Many approaches to the data analysis were considered. (In fact, two alternative methods are presented at the end of Chapter Five.) As a result, much discussion took place between this author and the employees at the division office of the bank with respect to the decision process followed. It was decided that the best approach would be a "pooled approach" where all the data (339 cases), from both the branch and division management levels, were considered to be part of one large sample. Hence, the outcome being modelled entails the approval or the decline of the loan application by either the branch or division manager. Thus, the sample is comprised of 252 approvals and 87 declines. As was described above, 61 of the 87 declines were obtained from 4 branches selected by Division Management. This oversampling of 4 particular branches insures the significance of the branch effect in any subsequent model building exercise. However, this fact does not diminish the significance of the "pooled approach" model. Firstly, the variable selection process is not disturbed. Secondly, the relationship between the decision process being modelled and the predictor variables are unaffected by this oversampling. Lastly, the use of validation techniques are not hampered by this result. Consequently, since the pooled approach best represents the decision process (see below), the oversampling of the branches will be tolerated and the limitations on the branch effect are so noted. Therefore, the majority of the analysis in this thesis, such as examination of residuals and validation techniques, pertains to this modelling plan.

In order to justify this approach, we must consider that the decisions at this bank (at the branch or the division offices) are made by management trained in the same procedures at the same division of one bank. It is reasonable to assume that all managers, regardless of level, should follow the same decision criteria. If this is the case, then it is tenable to group all of the declines into one outcome level (y=0, the no group) and the approvals into the other (y=1, the yes group).

A further justification to the pooled approach is that so few cases are rejected by Division Managers (less than ten percent - only 26 out of 278). This indicates that most of the bad applicants have been screened out by the branches. (Based on discussions with bank management, this branch rejection rate can be as high as 70 to 80 percent.) The Division Managers, then, for the most part, approve applicants which have proven their credit-worthiness to the branch managers. As a result, there seems to be a similar decision process in place at these two levels of management.

One objection to this pooled approach which could be raised is that the branch managers recommended 26 cases which were later declined at the division level. Does this factor contradict the hypothesis that all managers, regardless of level, follow the same decision process ?

In response, it should be emphasized that these 26 cases were recommended by the Branch Managers - not approved. It is possible and highly probable that a Branch Manager would recommend a higher percentage of cases than he would approve. Any borderline case could be recommended with the awareness that some individual higher up in the organization is responsible for the final decision. Hence, a "recommendation" and an "approval" are not the same. Given the authority and the responsibility of the "final say," the Branch Managers may very well have declined these 26 cases. A second explanation could be that the Division Managers correct the errors which have been made by the Branch Managers. These Branch Managers are less experienced than their division counterparts and are, therefore, more prone to make errors.

In short, the data from both samples (division and branch level) are grouped together to form one large sample. The main objective is to simulate the decision process of the loan officer (regardless of level) at the time of the loan application (OLA).

CHAPTER FOUR

STATISTICAL METHODOLOGY

The objective of this chapter is to build a credit scoring model for predicting the approval or the rejection of a small business loan application based on a number of predictor variables. As a direct result, a list of important predictor variables is generated which will be of interest both to lending institutions and to the small businesses.

4.1 Application of Statistical Techniques to Bank Data

Once the data were fully coded and a data file was created, they were analyzed using a number of statistical techniques. Initially, the problem of multicollinearity was investigated. An indepth analysis of the correlation matrix, as well as a principal components analysis, were performed in order to eliminate highly correlated predictor variables. A discussion of the results of this and all other analyses presented in this chapter is provided in Chapter Five. (For a more detailed discussion of estimation problems and multicollinearity, see, for example, Neter, Wasserman and Kutner [1985].)

Subsequent to this preliminary analysis, multivariate normal linear discriminant analysis was employed. This analysis determines the important variables involved in the evaluation of credit. The linear discriminant model was used by Doreen and Farhoomand (1983) in their study of the OLA (outcome of the loan application) for small business loans. Utilization of this model (and the subsequent results) provides a reference point for comparison to the other statistical techniques.

Very briefly, the objective of discriminant analysis is to construct a classification scheme whereby previously unclassified observations are assigned to groups. Group membership is determined on the basis of certain characteristics (predictor variables) possessed by the observation being classified (Eisenbeis and Avery, 1972).

The three basic assumptions of the multivariate normal linear discriminant model are:

(i) the groups being considered are distinct and identifiable;
(ii) the independent variables have a multivariate normal distribution; and
(iii) the group dispersion matrices are equal.

If these three assumptions hold, then linear discriminant analysis can be applied. If the assumptions do not hold, then further analysis should be conducted to determine the effects of the departure(s). In addition, other techniques should be considered as possible alternatives for the problem in question.

When the parameters of the discriminant model must be estimated from the data, a further condition on the sample observations is that they be randomly drawn from their respective groups. In other words, for this study, the "no" observations must be considered as a random sample from a population of no's and the "yes" observations must be considered a random sample from a population of yes's. Recall that all observations (within the limits specified in Chapter Three) are retained in the sample for both groups. Thus, if we consider these 48 branches (or this one division) as representative of a larger population of similar branches (perhaps of all divisions across the country for the one bank or of all banks across this division), then we have a random cluster sample. Strictly speaking then, the assumption of random sampling has not been met. In order to adjust for the sampling scheme, parameters associated with branches are included in the model used to analyze these data.

One particularly important assumption for linear discriminant analysis is the one pertaining to the distribution of the predictor variables. Firstly, the independent variables are often assumed to be multivariate normally distributed within groups. Surprisingly, there has been very little acknowledgement of this assumption in credit scoring research (Reichert, Cho and Wagner, 1983). Researchers proceed as if the normality assumption holds for all continuous independent variables (Eisenbeis, 1977). (See Ameniya [1981] for a discussion of the robustness of the discriminant model.)

Secondly, categorical independent variables may pose a problem. When the assumption of randomness holds, these variables inherently possess a multinomial, not a normal, distribution (Wiginton, 1980).

Thirdly, a model can contain a mixture of categorical and continuous predictor variables. One possible solution to this problem is to split the samples based on the values of the discrete variables and analyze the split samples separately (Chang and Afifi, 1974). If this is not possible, then other statistical techniques should be investigated.

These three points illustrate some of the possible shortcomings of linear discriminant analysis. One alternative to linear discriminant analysis when the outcome variable is binary (0,1) is a regression model on a transformed scale. The objective of this type of analysis is to model the value of p - the probability of the loan being accepted - rather than the (0,1) outcome variable. The reason for a transformation of scale is obvious. When predicting a probability (a value from zero to one), a linear regression model may not be suitable because the predicted values could extend to values greater than one or less than zero (outside the feasible region). On an appropriate transformed scale, predicted values could be restricted to the (0,1) range. Several possible transformations could be considered, such as the integrated normal (probit) transform or the angular (arc sine) transform (see Fienberg, 1980).

The model considered here is a regression model transformed to the logistic scale - logistic (or logit) regression. In general, the logit (p) is defined as the log (p/(1-p)), or the log-odds, where the value of p (a probability) is being estimated in order to predict the outcome of the (0,1) variable. For a full discussion of logit regression, see Bishop, Fienberg and Holland (1975), Fienberg (1980), Cox (1970), and Haberman (1978).

Let us consider, for the purpose of illustration, an example of a logit model for the credit scoring data:

$$\text{logit}\,(p_i) = b_0 + u_{1(i)} \qquad (4.1.1)$$

where we have the following representation:

p_i represents the probability of a loan being granted to any applicant where i represents the branch of the bank where the application was filed;

b_0 represents the overall constant term for the model;

$u_{1(i)}$ represents the effect of the branch of the bank.

(A more detailed model is examined in Chapter Five.)

In order to motivate the discussion for a third statistical technique we refer to the model given in equation (4.1.1). The $u_{1(i)}$ term represents the effect of the branch of the bank where the application was filed. There are many possible categories or branches. In fact, the data contain loan applications from a sample of 48 branches. However, we are interested in making inferences for all branches - not just those that were observed in the sample. As a result, we would like to be able to extend our inferences from this sample to the entire population of branches.

Secondly, there are branches for which little information is available. It would be useful to exchange or share information from other branches with more data in order to get better estimates for those with few observations.

Thirdly, incorporated into the effect of the branch are many hidden factors. For example, the actual decision criteria of the particular Senior Branch Manager, the type of industry in the region of the branch, the level of competition in the different industries, and any initial screening effect at the branch - these are all factors pertaining to the branch effect. Thus, it is important to study the variability of this branch effect.

One way of resolving all of these problems is to consider the effect of the branch to be random. We consider a population of branches, where the 48 in this particular division form a random sample. This assumption of randomness allows for the incorporation of an "empirical Bayes rule" for multi-parameter estimation. Briefly, branch effect estimates shrink toward an overall central estimate for the branches where the sample information is scarce. This shrinkage technique often leads to better estimates than those obtained by maximum likelihood (to be discussed in Section 4.2). (See Efron and Morris [1972,1973,1975] for a full discussion.)

The assumption of randomness applies only to the branch effect. Other effects that are included in the model are treated as fixed since there is no justification to treat them otherwise.

Consequently, a third technique, logit regression with random effects, will be used. Wong and Mason (1985) introduced a hierarchical logistic regression model with random slope coefficients and have applied it to a multi-nation survey concerning the use of birth control devices. The parameters of these random coefficients were estimated through a multilevel logit regression analysis involving discrete and continuous predictors. The study emphasizes the method of estimating these parameters.

In this dissertation, a similar model, incorporating some modifications, is applied to the credit scoring problem. The first modification is that a logit regression model with both fixed and random parameters (a mixed model) is considered. Secondly, although parameter estimation is of interest, the main emphasis of this research is on "prediction". Once again, we are primarily interested in the prediction of the outcome of the loan application (OLA).

Based on the data which have been collected, an initial logistic regression model with fixed and random effects is presented; this is for illustration purposes only:

$$\text{logit}\,(p_{ijkl}) = b_0 + u_{1(i)} + u_{2(j)} + u_{3(k)} + b_1 X_{ijkl} \qquad (4.1.2)$$

where $u_{1(i)} \sim N(0, \sigma^2)$.

p_{ijkl} represents the probability of a loan being granted to any applicant, where

i indicates the branch of the bank where the application was filed, $i = 1, \ldots, 48$;

j indicates the type of loan, $j = 1, \ldots, 4$;

k indicates whether the company is a new or not and the type of owner history, $k = 1, \ldots, 4$;

l indicates the observation in the ijk-th cell,

$l = 1, \ldots, n_{ijk}$; n_{ijk} = number of applications in the ijk-th cell.

For each loan application, we are attempting to estimate p_{ijkl}.

The b_0-term represents the overall constant for the model. The X_{ijkl} term denotes the amount of the loan.

The $u_{1(i)}$ term represents a branch effect, which is considered random and is regarded as having been drawn from a normal population with mean zero and variance σ^2. The variance term can be estimated using an EM algorithm (Dempster, Laird, and Rubin, 1977). This value indicates the amount of variation in the decision process from branch to branch. A large value of σ^2 could indicate that some important independent variables have been omitted.

The $u_{2(j)}$ term represents the type of loan being applied for and is considered to be a fixed effect (as described above). There were only three types of loans encountered - operating loans, term loans, and term loans under the Small Business Loan Application (SBLA) program (which is government guaranteed). We restrict our inferences to the levels of this variable observed in this study.

The $u_{3(k)}$ term indicates whether the firm is a new company, one already in existence, or one new to the bank. All the levels of this categorical variable were observed in the sample, and we are interested in inferences pertaining solely to these levels. Consequently, we treat this term as a fixed effect.

4.2 Estimation Procedures

Initially, we follow previous work in log-linear models by Leonard (1975) and Tomberlin (1982, 1987), and we incorporate Bayesian and empirical Bayesian methods to analyze the model (eq. 4.1.2). Under pure Bayesian philosophy, the prior distribution is specified - a value for the variance term is given for both the fixed and random effects. Under empirical Bayes, however, the variance is not specified, but is estimated from the data. This approach is quite viable in a setting such as this where there are many parameters involved. Thus, for the random effect terms, the variance is calculated from the observations; for the fixed effects, a flat prior is specified.

The next two sub-sections give the initial Bayes and empirical Bayes estimates for this model. Iterative algorithms are discussed because they are needed to obtain the final estimates.

4.2.1 Bayes Estimates

The conjugate prior for Bayesian analysis of Bernoulli data, where the parameter is p, is the beta distribution (see DeGroot, 1970, for example). However, this prior distribution does not lend itself well to exploiting the structure of a logit regression model. On the other hand, the multivariate normal distribution adapts readily to these linear ANOVA type models, and has received some attention. Thus we consider a multivariate normal prior distribution for the parameters in the model. (See Tomberlin, 1982, for similar applications based on the Poisson distribution; and Laird, 1978, and Leonard, 1975, for applications based on the multinomial.)

For notation purposes, let us consider a vector, **u** , which has as its components the values for (b_o, b_1, $u_{1(i)}$, $u_{2(j)}$, $u_{3(k)}$). In other words, **u** represents a vector of parameters which contains the slope parameters, as well as, all fixed and random effects parameters. Then, by specifying a flat prior on b_o, b_1, $u_{2(j)}$, and $u_{3(k)}$, and a normal prior on $u_{1(i)}$, we obtain the joint distribution of **y** and **u** as

$$p(y, u \mid X, \sigma^2) = p(y \mid u, X, \sigma^2)\ p(u \mid X, \sigma^2) \qquad (4.2.1)$$

$$= \prod_{ijkl} [p_{ijkl}(u)^{y_{ijkl}} \cdot (1-p_{ijkl}(u))^{1-y_{ijkl}}] \cdot \qquad (4.2.2)$$

$$[\frac{1}{\sqrt{2\pi}\ \sigma}]\ \exp\ [\frac{-1}{2} \sum_i \frac{u_{1(i)}^2}{\sigma^2}]$$

where y_{ijkl} = 1, or 0, denoting success or failure of the application for each loan, respectively,

y = the matrix containing these results,

u = the vector of parameters, the u-terms and the b-terms,

X = the matrix of predictor variables, and

$$p_{ijkl}(u) = \frac{\exp[b_0 + u_{1(i)} + u_{2(j)} + u_{3(k)} + b_1\ X_{ijkl}]}{1 + \exp\ [b_0 + u_{1(i)} + u_{2(j)} + u_{3(k)} + b_1\ X_{ijkl}]} \qquad (4.2.3)$$

Under traditional Bayesian analysis, with σ^2, the prior variance of $u_{1(i)}$, specified in advance, estimation would be based on the posterior distribution given the prior variance, σ^2, the observed results, y, and the predictor variables, **X**. Thus we would have

$$p(u \mid y, X, \sigma^2) = \frac{p(y, u \mid X, \sigma^2)}{p(y \mid X, \sigma^2)} \qquad (4.2.4)$$

Due to the mathematical intractability of the integration of the denominator, we follow the work of Leonard (1975) and Laird (1982) in similar log-linear models, and approximate $p(\mathbf{u} \mid y, X, \sigma^2)$ by a normal distribution - its mean is at the mode of $p(\mathbf{u} \mid y, X, \sigma^2)$, and its covariance matrix is obtained by equating the second derivative matrix of the log of the normal approximation evaluated at the mean to the log of $p(\mathbf{u} \mid y, X, \sigma^2)$ evaluated at its mode. Effectively, the log posterior distribution, $\ln [p(\mathbf{u} \mid y, X, \sigma^2)]$, is approximated by a quadratic in the **u**-terms in the neighborhood around the mode (Laird, 1975).

4.2.1.1 Calculation of the Mode of the Posterior Distribution

As a first step, we take the natural logarithm of the posterior distribution. We then take the partial derivatives (with respect to the **u** components), set them equal to zero and solve simultaneously. This yields the mode for the posterior distribution. This methodology allows us to ignore the denominator in (4.2.4) due to the fact that it does not depend on **u**. These are the resulting partial derivatives equated to zero:

$$\frac{\partial}{\partial b_o} \ln p(\mathbf{u} \quad y, X, \sigma^2) = \sum_{ijkl} y_{ijkl} - \sum_{ijkl} p_{ijkl}(u) = 0 \qquad (4.2.5)$$

$$\frac{\partial}{\partial u_{1(i)}} \ln p(\mathbf{u} \quad y, X, \sigma^2) = \sum_{jkl} y_{ijkl} - \sum_{jkl} p_{ijkl}(u) - u_{1(i)}/\sigma^2 = 0 \qquad (4.2.6)$$

$$\frac{\partial}{\partial u_{2(j)}} \ln p(u \;\; y, X, \sigma^2) = \sum_{ikl} y_{ijkl} - \sum_{ikl} p_{ijkl}(u) = 0 \qquad (4.2.7)$$

$$\frac{\partial}{\partial u_{3(k)}} \ln p(u \;\; y, X, \sigma^2) = \sum_{ijl} y_{ijkl} - \sum_{ijl} p_{ijkl}(u) = 0 \qquad (4.2.8)$$

$$\frac{\partial}{\partial b_1} \ln p(u \;\; y, X, \sigma^2) = \sum_{ijkl} y_{ijkl} X_{ijkl} - \sum_{ijkl} p_{ijkl}(u) X_{ijkl} = 0 \qquad (4.2.9)$$

$$\text{where } p_{ijkl}(u) = \frac{\exp\,[b_0 + u_{1(i)} + u_{2(j)} + u_{3(k)} + b_1 X_{ijkl}]}{1 + \exp\,[b_0 + u_{1(i)} + u_{2(j)} + u_{3(k)} + b_1 X_{ijkl}]}$$

and X_{ijkl} is the "amount of the loan application" value (AMT) for the lth applicant in the ijk-th cell.

When this system of simultaneous equations is solved, the solution, denoted by $u = u^*$, represents the Bayes estimates. It can be shown that if we had specified a flat prior for $u_{1(i)}$ (i.e., $\sigma^2 \rightarrow \infty$; the prior information is weak), then equations (4.2.6; 4.2.7; 4.2.8) would have implied equation (4.2.5) - the total number of predicted successes $[p_{ijkl}(u)]$ in the BRN, TY and NEW cells must equal the total number of observed successes $[y_{ijkl}]$ across all (ijk) cells. On the other hand, if σ^2 had been set to an extremely small value ($\sigma^2 \rightarrow 0$), this would then have had the effect of forcing the estimates of the $u_{1(i)}$ terms to equal a common mean value, zero, specified in the prior distribution (see Tomberlin, 1982 for a full discussion of this topic). The effect of specifying a prior value somewhere between these two extremes causes the parameter estimate to depend on both components - the data observed in the cells and the overall mean value of the branch effects.

The actual calculation of the Bayes estimates follows. The set of non-linear equations (equations (4.2.5) through (4.2.9) above) is solved simultaneously. A number of algorithms are available to perform this function. The one that was selected was subroutine (ZSPOW) from the International Mathematics and Statistics Library (IMSL, 1982) which employs a quasi-Newton method. (See Appendix B for the computer programs which were implemented to solve these equations.) The solution to these equations provides the Bayes estimates for the parameters.

4.2.1.2 Calculation of the Posterior Covariance Matrix

The posterior covariance matrix, Σ, is obtained by taking the negative of the inverse of the second derivative matrix of the log of the posterior distribution, equation (4.2.4), evaluated at the mode (i.e., at $\mathbf{u} = \mathbf{u}^*$). The inverse of this matrix, Σ^{-1}, is given in equation (4.2.10). With the inversion of this matrix, the calculation of the Bayes estimates is complete. We have approximated the true posterior distribution with a normal distribution having

INVERSE OF THE COVARIANCE MATRIX (Σ^{-1}) (4.2.10)

$$
\begin{array}{c|cccccccccccc}
 & b_0 & U_{1(1)} & U_{1(2)} \cdots & U_{1(I)} & U_{2(1)} & U_{2(2)} \cdots & U_{2(J)} & U_{3(1)} & U_{3(2)} \cdots & U_{3(k)} & b_1 \\
\hline
b_0 & \sum\limits_{ijkl} p(u)[1-p(u)] & & & & & & & & & & \\
U_{1(1)} & \sum\limits_{jkl} p(u)[1-p(u)] & \sum\limits_{jkl} p(u)[1-p(u)]+1/\sigma^2 & & & & & & & & & \\
U_{1(2)} & \sum\limits_{jkl} p(u)[1-p(u)] & 0 & \sum\limits_{jkl} p(u)[1-p(u)]+1/\sigma^2 & & & & & & & & \\
\vdots & \vdots & \vdots & & & & & & & & & \\
U_{1(I)} & \sum\limits_{jkl} p(u)[1-p(u)] & 0 & \cdots\ 0 & \sum\limits_{jkl} p(u)[1-p(u)]+1/\sigma^2 & & & & & & & \\
U_{2(1)} & \sum\limits_{ikl} p(u)[1-p(u)] & \sum\limits_{kl} p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{ikl} p(u)[1-p(u)] & & & & & & \\
U_{2(2)} & \vdots & \vdots & & & 0 & & & & & & \\
\vdots & & & & & \vdots & & & & & & \\
U_{2(J)} & \sum\limits_{ikl} p(u)[1-p(u)] & \sum\limits_{kl} p(u)[1-p(u)] & \cdots & \cdots & 0 & \cdots\ 0 & \sum\limits_{ikl} p(u)[1-p(u)] & & & & \\
U_{3(1)} & \sum\limits_{ijl} p(u)[1-p(u)] & \sum\limits_{jl} p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{il} p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{ijl} p(u)[1-p(u)] & & & \\
U_{3(2)} & \vdots & \vdots & & & \vdots & & & 0 & & & \\
\vdots & & & & & & & & \vdots & & & \\
U_{3(k)} & \sum\limits_{ijl} p(u)[1-p(u)] & \sum\limits_{jl} p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{il} p(u)[1-p(u)] & \cdots & \cdots & 0 & \cdots\ 0 & \sum\limits_{ijl} p(u)[1-p(u)] & \\
b_1 & \sum\limits_{ijkl} X_{ijkl}\, p(u)[1-p(u)] & \sum\limits_{jkl} X_{ijkl}\, p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{ikl} X_{ijkl}\, p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{ijl} X_{ijkl}\, p(u)[1-p(u)] & \cdots & \cdots & \sum\limits_{ijkl} X^2_{ijkl}\, p(u)[1-p(u)]
\end{array}
$$

mean and covariance matrix as calculated above.

Once again, a subroutine (LINV2F) from IMSL was used to invert the matrix. (This program is also listed in Appendix B.)

4.2.2 Empirical Bayes Estimates

In empirical Bayes estimation, the unknown variance component, σ^2, is estimated from the data. Unlike Bayesian analysis, the prior variance is not specified in advance, but rather information contained in the present data is used to estimate σ^2. In this study, we have a multiple parameter estimation problem that lends itself well to empirical Bayes treatment. Empirical Bayes analysis forces the branch effects estimates to shrink toward a common central estimate in the branches where the sample information is not strong. Here, an EM algorithm is used to obtain an estimate for this prior variance.

The EM algorithm, described by Dempster, Laird and Rubin (1977), provides an iterative procedure for maximum likelihood estimation that is applicable in a general "incomplete data" setting. They present empirical Bayes estimation as an example. The term EM stands for two steps in the algorithm - the E step (expectation) and the M step (maximization). In order to implement the EM approach here, an initial value for the prior variance, σ^2, is specified; the y values are treated as observed values; and the u-terms are treated as missing values (see equation 4.2.4).

More specifically, each iteration of the EM algorithm requires a Bayesian

PLOT OF LIKELIHOOD FUNCTION

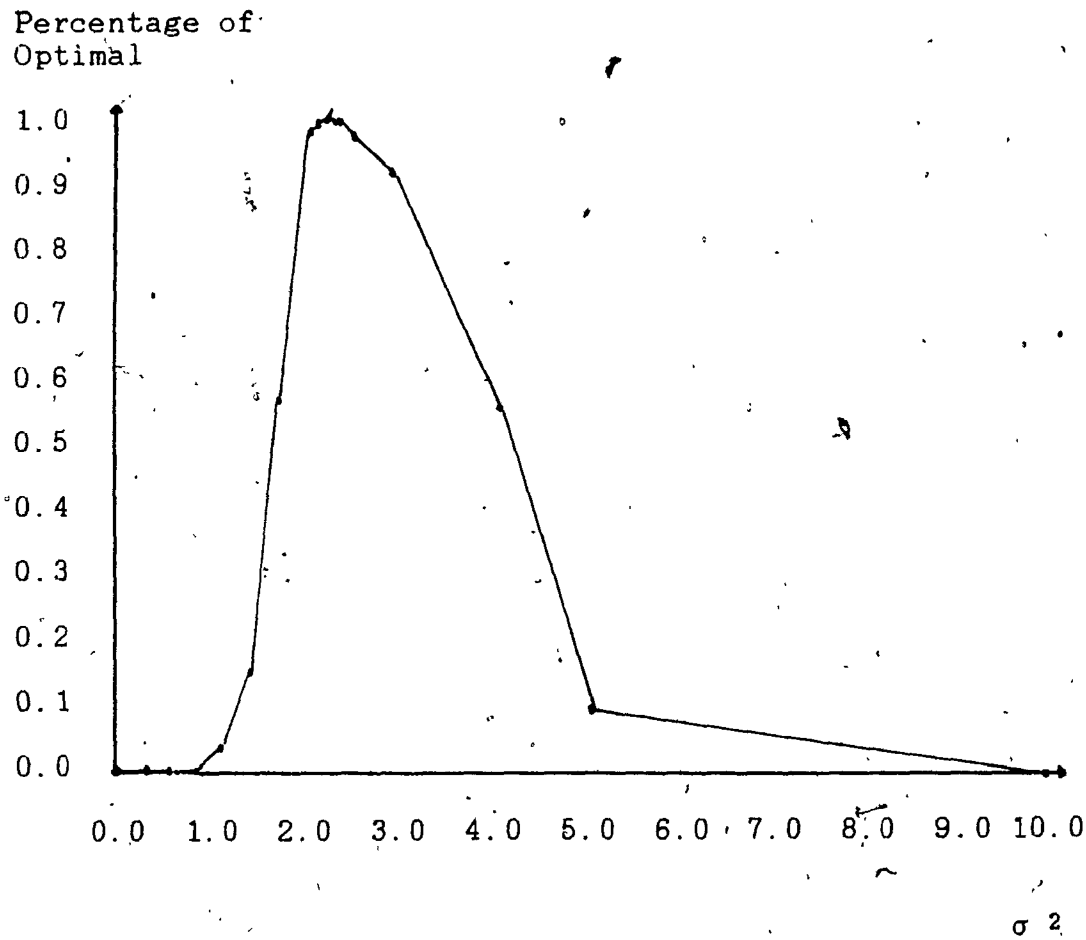


Figure 4.2.1

analysis as described in the previous subsection. The E-step of a single iteration corresponds to deriving the expected value of the log-likelihood [ln p(y, **u** | X,σ^2)] with σ^2 set at a starting value of σ^2(o) where the u-terms are considered random. At the M-step, the value for σ^2 is estimated by maximizing the likelihood function of the sufficient statistics based on the posterior distribution obtained from the Bayesian analysis. (This maximization step is simplified due to the nature of the family of regular exponential distributions and the corresponding sufficient statistics [Laird, 1978].) The algorithm then iterates back to the E-step, using the new value for σ^2 for deriving the expected value of the log-likelihood [ln p(y | X, σ^2)] - this updates the u-terms. The algorithm continues until it converges and the empirical Bayes estimates are obtained. This results in a parameter estimate for σ^2, as well as, empirical Bayes estimates for the coefficients in the model - the u-terms (both the fixed and random effects terms) and the b-coefficients.

4.2.3 Illustration of the Likelihood Function

An interesting analysis which can be performed is an investigation of the shape of the likelihood function [ln p(y | X, σ^2)]. Although, this is not a necessary exercise when using the EM algorithm, it gives an indication of the accuracy of the estimates of the prior variance. For example, when the shape of this function is peaked and/or concentrated, this would indicate a more precise estimate. The value of this likelihood function should be maximized when σ^2 is set at the value determined by the EM algorithm.

The procedure of estimating the likelihood function is as follows: Various values for σ^2 are employed as starting points. For each value the system of simultaneous equations (eg. 4.2.5 through 4.2.9) is solved and the Bayesian estimates are obtained. These estimates are then put into the equation and the value of the likelihood function, [ln $p(y \mid X, \sigma^2)$], is calculated; the value relates to the corresponding starting value of σ^2. When enough points are determined, a plot can be developed (Tomberlin, 1982).

At this stage, using the preliminary model only (equation 4.1.2), an initial plot can be based on the prior variance, σ^2, in order to illustrate the data. The plot is given in Figure 4.2.1 and the listing of points is given in the table below.

σ^2	Values of ln $p(y \mid X, \sigma^2)$ as a percentage of the optimal point
0.000	0.0000
0.250	0.0000
0.500	0.0000
0.750	0.0000
1.000	0.0439
1.250	0.1473
1.500	0.5637
1.750	0.9828
1.900	0.9991
2.017	1.0000
2.200	0.9923
2.500	0.9415
3.000	0.8932
4.000	0.5611
5.000	0.0713
10.000	0.0000

This procedure is repeated in Section 5.8 of this paper pertaining to the "separation" models where there are two prior variances, σ_1^2 and σ_2^2. Grid points and a three dimensional plot are provided to describe the shape of the likelihood function for this application.

4.3 Model Building and Variable Selection

This section concentrates on diagnostic methods pertaining to model building. Once a subset of independent variables was produced (eliminating highly correlated predictors), different models were examined for linear discriminant analysis (LDA), and for logit regression with fixed effects (LRFE).

More specifically, using a statistical package for stepwise discriminant analysis (BMDP7M-83), different LDA models were considered. A more detailed discussion on this topic is provided in Chapter Five.

A similar approach was taken for LRFE. Again using a stepwise package (BMDPLR-83), the "best" model was produced. The set of predictors chosen by this procedure was used in the LRME (logit regression with mixed - fixed and random - effects) model. The execution of the LRME is a very costly and time consuming process (using the EM algorithm). Consequently, stepwise methods were not applied to the LRME model. It is possible that a more suitable model for LRME could have been produced if a more extensive and cost efficient variable selection methodology was available. However, this is left for future research.

Once the "best" models for LDA, LRFE, and LRME were produced, an examination of the goodness of fit took place using residual analysis. Although residual analysis has been much discussed in the recent statistical literature, at this time, there is no work comparing residual analyses for all three of the statistical methods discussed here - LDA, LRFE, and LRME.

4.4 Comparison of the Statistical Methods

The main emphasis of this research is on the prediction of the outcome of the loan application (OLA). As such, much effort is spent on classification results. Analysis is divided into three parts:

(a) Classification of the original data set

This involves testing each of the "best" models on the original data set and comparing the results across the three methods.

(b) Classification of a holdout data set

Part (a) evaluates the three methods' effectiveness as classification procedures. In part (b), comparison is based on classifying a new data set (prediction). Following the suggestions of many authors (for example, see Reichert, Cho and Wagner [1983]), we split the original data set into two subsets. The "best" models for LDA, LRFE, and LRME were fit using the first data subset (the training sample). These fitted models were then tested on the second data subset (the holdout sample). The classification results for the holdout sample indicate how well the models can predict new cases.

Here, we consider two different holdout sample sizes - a fifty percent sample (of the original data set) and a twenty percent sample. This "holdout procedure" was employed ten times - five times using the 50 percent split and five times using the 20 percent split.

c) Estimation of true prediction error using the bootstrap method

Another method for measuring classification effectiveness is the estimation of the true prediction error rate (see Efron, 1982,1983; Gong, 1986). The true prediction error rate is defined as the rate of error found when the true classification model predicts the entire population of cases. However, this value cannot be determined (the population is not available), and, as such, can only be estimated.

Before we proceed with this estimation, a couple of terms should be introduced. The apparent error rate is defined as the rate of error arrived at when the model created from the sample is used to predict that very same sample (usually, the original data set). The expected excess error rate (EEE) is defined as the true error rate minus the apparent error rate.

In order to estimate the true error rate, we must estimate the value of the EEE. A number of methods for estimating EEE have been suggested with the bootstrap procedure being the method favoured most by Efron (Efron, 1982).

The bootstrap method has been used for estimation in a number of statistical settings. This method is especially useful for studying sampling distributions. In general, the methodology is as follows: Random samples (we will refer to these as the "resamples") are drawn with replacement from the observed values in the original sample. Sample estimates can then be obtained using the resamples, providing an indication of sample to sample variability. The data in the original sample are treated as the population for estimation purposes (see Efron, 1982, for a full discussion).

The bootstrap estimate of the EEE is determined in the following manner. First, a random sample, equal to the size of the original data set is selected (with replacement) - call this the "resample". Then the "best" models for LDA, LRFE, and LRME are fit using the data in the resample. The classification error rates determined here are the "estimates" of the apparent error rates. These same models are then tested on the original data set (observed values). The prediction error rates on the observed values are the "estimates" of the true error rates. The differences between these two error rates are the estimates of the EEE rates.

This bootstrap procedure can be replicated many times, and an overall estimate of the EEE rate can be determined for the three methods. Here, the number of replications is limited to twenty due to the slow convergence and resulting high cost of the EM algorithm.

The final step requires the addition of the estimate of the EEE to the apparent error rate (classification of the "best" models on the original data set) and this produces an estimate of the true error rate for the LDA, LRFE, and LRME models.

4.5 Missing Data Estimation

A common problem with empirical analysis is missing data. Missing data may occur due to sampling deficiencies or to an absence of information on certain variables encountered during data capture. A number of techniques are available to handle this problem. One such solution is to simply remove the cases where the data are not complete and perform the analysis on the rest of the data set. (This method is used by the BMDP package.) This approach was applied to the data set in this analysis to determine the best models for discriminant analysis and logit regression. For this data set, 56 cases (out of 339) were removed for the LRFE, and none were removed for the LDA model (due to the fact that missing values only occurred for the total net worth or the TNW variable).

An alternative approach is to estimate the missing values and then use all the cases (the entire data set) to fit the different models. This technique is employed using the EM algorithm to determine the estimates for the missing values of total net worth (TNW). This is the same EM algorithm that is discussed in Section 4.2. In fact, the EM algorithm was initially developed to handle the task of estimating missing observations. Basically, the relationship of the other variables observed and the variable for which values are missing are analyzed. Information on the cases where all the data are present is pooled to help estimate values for the cases where the data are missing. More specifically, where there are missing values, we substitute for these values their expectations given current parameter values (E-step), and then re-estimate the parameters using a simple least squares algorithm (M-step) (Dempster, Laird and Rubin, 1977). This algorithm iterates back and forth until the estimates exhibit no important change.

It is possible to estimate the missing values <u>and</u> to estimate the parameters for the logit regression model with mixed effects simultaneously using the EM algorithm. However, this procedure was not followed here, primarily due to the complexity of the task. We chose to handle the two tasks (estimation of missing data and calculation of parameter estimates) separately and reduce the programming and computational burden.

The results for this and all preceding analyses are presented in Chapter Five.

CHAPTER FIVE

DATA ANALYSIS

In this chapter, the results and interpretations of the data analysis are presented. All tables and figures referred to appear in Appendix A. As well, some of these tables and figures have been reproduced and placed in the chapter in order to provide the reader with ease of reference.

As was discussed in Section 3.5, we collapsed the data from the two samples (the Division Managers and the Branch Managers) to form one overall sample. The objective, then, is to model the decision process of the loan officer, regardless of management level, with respect to the outcome of the loan application. The majority of this chapter pertains to this model, and it includes extensive analysis of residuals and validation techniques.

After this initial application, we examine two alternative methods for modelling the bank data - a "separation" model and a "multi-model" approach.

5.1 Results of Model Building - Variable Selection

As we documented in the previous chapter, one problem which can arise in regression analysis is multicollinearity. The use of correlated predictors can lead to difficulties in ascertaining the relative importance of different independent variables and in determining the magnitude of the effect these variables have on the dependent variable. However, it should be pointed out, that, multicollinearity does not adversely affect the predictive model. Thus if our sole aim is to create the best predictive model possible, we should not be concerned about correlated predictors. In this thesis, by employing the appropriate variables, we want both to predict and to create the best possible model; consequently, correlated predictors are removed as much as possible. The first step was to identify the correlation, if any, between the predictor variables. (See Neter, Wasserman, and Kutner, 1985, for a full discussion.)

In this data set, we consider approximately twenty predictor variables. A correlation matrix was calculated for all the variables in the full sample. Many variables were found to be highly correlated. For example, amount of the loan (AMT) and total credit position (CRE) were highly correlated ($r = 0.867$). This is as expected, since, for most of these small businesses, this loan application represents the major component of their external financing. Other variables which proved to be highly correlated were sales of the firm (SA), gross profit (GP), and total income after expenses (INC). We expect, that, for most firms, these three variables will be highly correlated since they all originate from the same portion of the Income Statement resulting in a common scale effect. (See Table A-1 in the appendix for the full correlation matrix.)

Functional relationships between variables were also investigated. For example, AMT/TNW and AMT/SEC were used as potential predictors in the model. Interactions with the branch effect, BRN*AMT and BRN*TNW, were also used. These terms were examined so that the proper functional form of the model could be obtained.

Once the highly correlated predictors were identified, a second procedure, principal components analysis, was performed (see Table A-2). The purpose of this analysis was to cluster the independent variables and then select the most representative variables within each cluster for use in the initial model. This was done in order to reduce multicollinearity and produce the best possible model.

Below is a list of the variables that were used as a starting set in the stepwise procedures (discriminant analysis and logit regression):

OUT	BRN	SEC
AMT	BRN*AMT	PUR
NEW	BRN*TNW	WC
TY	AMT/TNW	SA
TY*NEW	TNW	

5.1.1 The Linear Discriminant Model

Subsequent to the use of these two procedures, namely, analysis of the correlation matrix and principal components analyses, stepwise discriminant analysis was performed. A routine in the Biomedical computer programs -P series, BMDP7M (BMDP, 1983), was implemented. In the initial step, the variable with the highest "F-to-enter" is entered into the discriminant function. This is the variable that discriminates the best between groups. This "F-to-enter" statistic corresponds to the F statistic computed from a one-way analysis of variance on the variable for the groups used in the analysis. The degrees of freedom are (g-1) in the numerator and

$\sum_{k=1}^{g} (N_k - 1)$ in the denominator where k is the number of the group and N_k is

the number of observations in each group. At each subsequent step, an "F-to-enter" value is calculated for the independent variables absent from the equation. This F statistic is computed from a one-way analysis of covariance where the covariates are the previously entered variables. This process continues until such time that the "F-to-enter" statistic is not significant (i.e., the variable makes no significant contribution to the discriminant function). The "F-to-enter" cut-off point was 3.00, corresponding approximately to an F statistic with 1 degree of freedom in the numerator and a relatively large number of degrees of freedom in the denominator (> 120) with the level of significance between 0.05 and 0.10. This stepwise procedure also allows for any variable to be removed from the model if it is greater than the "F-to-remove" value. This "F-to-remove" value used was 2.996.

The stepwise results (the summary table) for the "best" model are given in Table A-3. The listing of the parameter estimates are under the heading LDA and are given in Table A-4 (and reproduced below). We can represent the "best" model as the following linear discriminant function:

$$Z_{ijkl} = b_0 + BRN_i + TY_j + NEW_k + TY_j * NEW_k + b_1 BRN_i*AMT \tag{5.1.1}$$

where Z_{ijkl} is the discriminant score which is compared to a predetermined cut-off point;

BRN is the branch of the bank;

BRN*AMT is the interaction of branch and amount of loan;

TY is the type of loan;

NEW represents the experience of the applicant; and

TY*NEW is the interaction between TY and NEW.

This model was produced by a stepwise analysis. It is possible that an equally or more efficient model could be produced through an "all-possible regressions" analysis. Thus, this may not be the "best" model. However, the stepwise approach offers good results and economizes on computational expense (Neter, Wasserman, and Kutner, 1985).

The variables which were found to be significant can be justified. The first variable included in the model is the BRN variable. The significance of this variable is due to the sampling procedure and the oversampling of the declines, as mentioned above.

There are other factors included in this branch effect which are related to the BRN variable which should be discussed. As stated in Chapter Three, a number of great restrictions were placed on the access to bank information.

TABLE A-4

		BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
		LDA*	LDA	LRFE	LRHE	LRFE	LRHE
CONSTANT		-2.3152	-1.25716	4.9269	1.9548	4.8441	2.2070
TY	(1)	.5394	0	0	0	0	0
	(2)	1.3097	-.93754	1.2091	.4381	1.1254	.3738
	(3)	2.7312	0	-.3274	.8109	-.5972	.3040
	(4)	0	0	1.4260	1.2590	1.3519	.9798
NEW	(1)	.4932	0	0	0	0	0
	(2)	1.2285	.5505	-2.3281	-1.0363	-1.9930	-1.0269
	(3)	.55154	0	2.9120	1.185	1.9614	1.2336
	(4)	0	0	.7524	.5617	.64681	.51925
BRN	(1)	.1765	0	0	.8210	0	.6055
	(2)	.9667	0	-1.2619	-.2556	-1.1670	-.1949
	(3)	1.3874	0	**	.0056	**	.0704
	(4)	3.0836	2.5004	-5.4090	-2.5364	-4.9235	-2.1782
	(5)	.9186	0	3.5395	.5996	3.7791	.4764
	(6)	-.1272	0	-2.0780	.32192	-2.1015	.22128
	(7)	.0000	0	**	.0007	**	.0233
	(8)	.0000	0	**	.0783	**	.0989
	(9)	.0000	0	**	.0453	**	.0007
	(10)	.5055	0	**	.6800	**	.554
	(11)	2.8073	2.34511	-5.0604	-2.345	-5.3829	-2.5259
	(12)	1.1686	0	**	.0423	**	.4329
	(13)	.0000	0	**	.2307	**	.2556
	(14)	.7855	0	**	.2584	**	.3578
	(15)	.1557	0	**	1.0276	**	.8447
	(16)	.0000	0	**	.0895	**	.1019
	(17)	1.1227	0	**	.1995	**	.3051
	(18)	1.0978	0	**	.8902	**	.7522
	(19)	4.1249	2.9247	-5.4316	-2.322	-5.0593	-2.3184
	(20)	.0000	0	**	.1202	**	.1023
	(21)	.6371	0	-1.6921	.18347	-1.3838	.3589
	(22)	.0000	0	**	.5268	**	.4171
	(23)	.0000	0	**	.1108	**	.09000
	(24)	.3586	0	**	1.1471	**	1.0351
	(25)	.0000	0	**	.0565	**	.0035

		BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
		EDA*	LDA	LRFE	LRHE	LRFE	LRHE
BRN	(26)	-.4832	0	∞	.6665	∞	.7156
	(27)	.1456	0	-1.7335	.0504	-1.4387	.1348
	(28)	.0000	0	∞	.4449	∞	.3748
	(29)	.0000	0	∞	.0763	∞	.0029
	(30)	-.9044	0	-1.3677	-.6744	-2.2225	-.5680
	(31)	.4851	0	∞	.8318	∞	.7940
	(32)	-1.6012	0	∞	.1279	∞	.3726
	(33)	.4845	0	48.497	-.8153	48.422	-.7168
	(34)	6.6881	6.11798	-19.881	-.6022	-20.049	-.6539
	(35)	1.3296	0	∞	.2321	∞	.2083
	(36)	1.1094	0	7.8243	.7798	-7.4096	.1434
	(37)	-.6767	0	-11.265	-.2718	-3.8630	-.0740
	(38)	.0000	0	-∞	-.9312	-∞	-.9428
	(39)	.0000	0	∞	.4076	∞	.3181
	(40)	.5553	0	-1.1681	.0474	-.96126	.2179
	(41)	.6309	0	∞	.6816	∞	.6896
	(42)	-1.32383	0	59.468	-.2756	49.458	-.3040
	(43)	.4325	0	∞	.5147	∞	.4720
	(44)	1.4828	0	∞	.3098	∞	.3525
	(45)	.6240	0	∞	.1875	∞	.1544
	(46)	-1.1373	0	213.99	-.4274	199.56	-.4166
	(47)	1.9594	0	-2.7595	-.6926	-2.8093	-.8304
	(48)	0	0	∞	.1694	∞	.4336
BRN*TNW	(1)			0	.000033	0	.000042
	(2)			.000469	.000505	-.000618	.00048
	(3)			—	.000032	—	.00001
	(5)			.03731	.00016	.03448	.0001123
	(6)			.00821	.000989	.00777	.000775
	(10)			.04895	.000152	.0447	.000108
	(7)			—	.000451	—	.00011
	(8)			—	.001007	—	.00009
	(9)			—	.000098	—	.000775
	(11)			.0140	.00349	.0157	.00325
	(12)			.0995	.00000204	.07174	.000053
	(14)			.0629	.0000495	.0271	.000045

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRHE	LRFE	LRME
BRN*TNW (13)			—	.006801	—	.00015
(15)			- .00801	.00022	- .00760	.000166
(17)			- .16021	.0000063	- .00994	.000041
(16)			—	.000367	—	.00072
(18)			- .006898	.000206	- .006127	.00016
(19)			.00005738	.0000264	.0002518	.000048
(21)			.00322	.000756	.00296	.00010
(20)			—	.00008	—	.00001
(22)			.02047	.000347	.0162	.000244
(24)			.00567	.000453	.0059	.000354
(23)			—	.00075	—	.00083
(26)			.00432	.000251	.00444	.000375
(25)			—	.000277	—	.000036
(27)			- .000338	.0001177	- .0003711	- .00000957
(28)			.0113	.000836	- .00752	.00062
(30)			- .0219	- .0002815	- .0106	- .00019
(29)			—	.0043	—	.0044
(31)			+ .00388	.00010	.00434	.000309
(32)			.00949	.000083	- .0116	.000087
(33)			- .81095	- .000227	- .80668	- .0001844
(34)			.28359	.0000799	.28394	.0000504
(35)			- .00194	.000055[illegible]	- .00250	.000039
(36)			- .0151	.000039	1.3944	.00010
(37)			.213	.0002[illegible]	.0548	.000156
(38)			- .6487	- .0000149	.967	- .00001286
(39)			.09165	.0000547	.0918	.0000378
(40)			.3958	.000394	.38677	.000332
(41)			- .00355	.000169	- .00323	.000119
(42)			-1.3828	- .0000778	-1.165	- .000081
(43)			.00675	.000043	.01009	.00[illegible]
(44)			.17018	.0000163	.15057	.00[illegible]
(45)			.00138	.0000251	0	.000[illegible]
(46)			-1.9887	- .000182	- 1.8532	- .0001[illegible]
(47)			- .00424	- .0011177	- .00424	- .00100
(48)			- .15717	.0000106	.0203	.000075
(4)			- .00518	- .000448	.004518	- .00069

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
NEW*TY (1,1)	-.4967	0				
(1,2)	-1.2596	0				
(1,3)	-2.3042	0				
(1,4)	.0000	0				
(2,1)	.0000	0.66337				
(2,2)	.0000	-0.67239				
(2,3)	.2396	0				
(2,4)	-2.1615	0				
(3,1)	-3.4704	0				
(3,2)	.0000	0				
(3,3)	.0000	0				
(3,4)	.0000	0				
(4,1)	-1.6621	0				
(4,2)	-2.1082	0				
(4,3)	-2.4964	0				
(4,4)	0	0				
AMT			- .003572	- .001497	- .003791	- .001381
AMT/TNW			.0013773	.001559	.006566	.002583
BRN*AMT (1)	.0010	0				
(2)	.0000	0				
(3)	.0000	0				
(4)	-.0001	0				
(5)	-.0051	0				
(6)	.0032	0				
(7)	.0255	0				
(8)	.0134	0				
(9)	.0294	0				
(10)	-.0026	0				
(11)	-.0011	0				
(12)	-.0376	0				
(13)	.0120	0				
(14)	.0144	0				
(15)	-.0002	0				
(16)	-.0434	0				
(17)	-.0141	0				
(18)	-.0035	0				

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
BRN*AMT (19)	-.0024	0				
(20)	.0060	0				
(21)	.0007	0				
(22)	-.0022	0				
(23)	-.1734	0				
(24)	-.0008	0				
(25)	-.0510	0				
(26)	.0023	0				
(27)	.0038	.00375				
(28)	-.0015	0				
(29)	.0201	0				
(30)	.0113	.00611				
(31)	-.0004	0				
(32)	-.0734	0				
(33)	.0126	.0122				
(34)	-.0539	-.05374				
(35)	-.0385	0				
(36)	-.0011	0				
(37)	.0159	0				
(38)	.2474	.22925				
(39)	-.0347	0				
(40)	.00156	0		$\sigma_1^2 = 1.8928$		$\sigma_1^2 = 1.6979$
(41)	.0000	0				
(42)	.0301	.01592				
(43)	.0007	0				
(44)	-.0247	0				
(45)	.0000	0		$\sigma_2^2 = .000004164$		$\sigma_2^2 = .00000331$
(46)	.0124	.00506				
(47)	.0059	.01224				
(48)	0	0				

NOTE:

TY (1) - operating loan
(2) - loan for equipment
(3) - loan for land and/or building
(4) - some combination of all three

NEW (1) - new firm with bank history
(2) - new firm with no bank history
(3) - existing firm with bank history
(4) - exisitng firm with no bank history

As a result, certain data were not available, and information concerning these different variables had to be determined through analysis of other variables. For example, there was no information collected concerning the type of industry pertaining to the customer. The commercial banking centres, however, are set up regionally, and certain information about the type of industry in a particular region may appear in the branch effect. Unfortunately, all these factors make an interpretation of the branch effect much more difficult. In addition, this logic may be extended to level of competition or any other factors which are regionally dependent.

It should be noted that the 48 different branches were entered into the discriminant function using dummy (or binary) independent variables. These variables were not entered in a block, and, as a result, many of the effects of the branches were not found to be significant. (Their coefficients are zero in the equation - see Table A-4 under the heading LDA. See Section 5.1.2 for an elaboration on this point.)

Let us examine the effect of Branch 4, for example, where the coefficient is 2.5004. The effect of this number is to decrease the likelihood of a loan being accepted; all other factors remain constant. (The cut-off of the discriminant score is +1.20; any score less than that indicates that the loan is approved.) Thus, for the sake of comparison, a positive value of approximately 2.5, with all other variables at their means, will cause the score to move from 0.837 - an approved application for a branch with effect zero, for example Branch 32 - to a score of 3.337 - a rejected application at Branch 4. This may suggest that more than an average number of loans are rejected at this latter branch. (See section 5.8 for a discussion of the branch effect and its relation to the sampling design.)

Second, the amount of the loan (AMT) variable is significant in the model. A relatively small firm may be financially sound enough to incur a debt of $10,000. However, this same firm could be considered a high risk if the loan application was for the larger amount of $500,000. There is an obvious relationship between the amount of the loan extended and the ability of the firm in question to repay that particular amount.

In the linear discriminant model, the amount of the loan is significant in its interaction with the branch level. This means that not only does the amount play a decisive role in the assessment of creditworthiness, but also the importance of this role varies depending on the branch. More specifically, the different branch managers evaluate the amount of loan variable with varying degrees of emphasis. In some branches, the manager will emphasize the requested amount and in other branches, the manager will consider the amount variable to have less bearing on the success of the candidate. It should be noted that the interaction term is significant in only a few of the non-oversampled branches which possess few declines. In the fixed effects model, when there are so few data points within the branch, the fixed effects model becomes tenuous. This situation is handled better within the framework of the random effects model, as is discussed in Section 5.1.4.

The third variable, type of loan (TY), is subject to similar factors.

Operating loans and term loans are used for different purposes, and, generally, the requirements for collateral are quite varied. For example, when granting a loan to a firm for the purposes of daily operations, credit managers may be concerned primarily with cash flow. On the other hand, when considering a term loan, bank managers may concentrate more on the asset being purchased or renovated and less on cash flow. This shift in concentration may be the main reason why the TY variable has been determined to be significant by the discriminant analysis. As an example, the coefficient for the second level (the second level represents a loan which is granted for the purchase or renovation of equipment) of the TY variable is -0.9375. Since this value is negative, it will increase the possibility of the loan being approved in accordance with the cut-off score and the acceptance region discussed above (the values for all other variables remaining constant). This type of loan is, generally, highly secured (the asset itself is used as collateral). Logically, this type of loan should have a high rate of approval in comparison to other types of loans.

Fourth, the NEW variable is included in the model. Once again, this variable contains information on the length of operation of the firm as well as the bank's past experience with the principal owners (managers) of the organization. The fact that this predictor is statistically significant might reflect that the bank managers involved in the decision process consider this variable to be quite important. If a firm is established in the marketplace and the bank enjoys a familiar and profitable relationship with the principal(s) of the firm, then credit managers will treat this customer in a favourable manner. However, if a firm is just getting started and the bank has had no previous experience with the people running the operation, then the bank may handle these customers in a manner different from that described for the customer above. This hypothesis is supported by the coefficients of these variables in the model. The coefficient for the case of a new firm with no experience is a positive number (+0.5505). This increases the score, and, consequently, increases the possibility of loan rejection.

The last variable included in the model, the interaction of the type of loan (TY) and the experience level of the customer (NEW), implies that the TY effect varies in accordance with the experience of the firm. For example, an operating loan for a new customer may be considered a high risk when compared to the same type of loan for an existing operation. One explanation for this difference may be that an existing firm has a past record indicating actual cash flow over similar periods of activity. Whereas, the calculations of cash flow activity for a new company are from pro-forma statements, which are only estimates of future activity. Again, the behaviour of the bank managers is supported by the values of the coefficients in the model. For a new firm with no experience (NEW - level 2), the effect is positive for an operating loan (TY - level 1), 0.66337. The financial statement information is obtained from pro-forma statements. However, for this same category of firm (NEW -level 2), a loan for equipment (TY - level 2) results in a negative effect, -0.67239. Once again, the collateral to secure a loan makes the extension of credit less risky, even to a new firm just starting operations.

5.1.2 The Adjusted Model (LDA*)

As was described briefly in the previous section, the categorical variables in the model (TY, BRN, NEW) were entered using binary (dummy) variables. This permitted only a subset of the levels of the variables to be entered. Initially, the variables were not entered as a block, and, consequently, all levels could not be considered simultaneously for each variable. In order to compare the LDA model with the logit regression models (to follow) where all levels of the categorical variables were entered together, the LDA model was run with the variables entering as a block. Therefore, the LDA* title represents the linear discriminant analysis (as discussed in Section 5.1.1) with the effects for categorical variables entered at the same time.

In addition, the adjusted discriminant model was run on only 283 cases (instead of 339) in order to match the number of cases used in logit regression. The logit regression uses less cases because there are observations missing for one of the variables (TNW). The identical data set allows for an equal comparison across methods.

The parameter estimates are listed in Table A-4 under the heading LDA*. A comparison of the two discriminant methods is difficult due to the different sizes of the data sets. It should be noted that the estimates are very similar in magnitude and direction.

5.1.3 The Logistic Regression Model

The next step in the research was to apply stepwise logit regression to the data. The criteria used to evaluate the different proposed models with selected variables were the log-likelihood statistics and chi-square values.

In more detail, at each step the set of coefficients, **b**, are estimated using the value that maximizes the likelihood function. After estimating **b**, a decision based on the log of the ratios of the maximized likelihood function (L) is made about whether to enter or remove any term in the next step

$$\chi^2 = 2\left|\log(L(\mathbf{b}\ \text{current})/L(\mathbf{b}\ \text{candidate}))\right| \qquad (5.1.2).$$

(See BMDPLR [BMDP-83] for a discussion.)

After the approximate chi-square values are obtained, the tail area probabilities are computed; the term with the largest p-value is removed if it is larger than the specified "remove" limit (a value of 0.10). If no term has a p-value larger than this limit, the term with the smallest p-value is entered if its p-value is less than the "enter" limit. The p-value 0.05 was used; it is similar to the cut-off employed for the stepwise discriminant analysis.

The stepwise results (refer to the summary table) for the "best" model are given in Table A-5. Once again, the parameter estimates are given in Table A-4 under the heading LRFE. As elaborated upon in Chapter Three in the discussion of the raw data, 31 branches had either all approvals or all declines. During the fitting of the fixed effects logit model, this causes the branch effect estimate to tend toward infinity. Thus, the branches where there were either all approvals or no approvals, the branch estimate is given as $+\infty$ or $-\infty$, respectively. When we are predicting a yes from a branch with all approvals, for example, the probability of a loan being granted will equal 1.0.

An alternative approach would be to remove these cases from the sample; the remaining information concerning other independent variables would then only be obtained from branches where both declines and approvals exist. However, this would result in a sample of only 17 branches and 246 cases, and the ability of the fixed effects model in prediction (holdout samples) would be greatly affected. Further, should we consider the entire sample of 48 branches and should any of the cases from these 31 branches are split into the training and the holdout samples, all cases in the holdout sample will be classified correctly. This gives an upper bound rather than a true indication of the effectiveness of the fixed effects model - thus the true effectiveness is disguised here as well. This is a problem inherent in the fixed effects model- not the data; consequently, we did not remove these cases from the data. Instead, the data set is left as is, and we note that prediction effectiveness may be upwardly biased. (As is shown in Section 5.1.4, the random effects model does not experience this problem.)

The "best" model can be represented as follows:

$$\text{logit}(p) = b_0 + BRN_i + TY_j + NEW_k + BRN_i{*}TNW + b_1 * AMT + b_2 * AMT/TNW \quad (5.1.3)$$

where p is the probability of a loan being approved;

BRN_i is the branch of the bank;

BRN_i*TNW is the interaction between branch and the total net worth of the firm;

TY_j is the type of loan;

NEW_k is the experience of the applicant;

AMT is the amount of the loan; and

AMT/TNW is the amount of loan divided by total net worth.

Many of the variables in this model are the same as the variables that were significant in the discriminant model. Although the coefficients of these variables (such as BRN and TY) are different for the various models (LDA versus LRFE), we have previously justified their placement in the model from a business (financial) perspective. Below, we concentrate on the new variables which were found to be significant under the logit regression. It should be remarked, that, in the discriminant model, both the interaction of TY and NEW and the interaction of BRN and AMT were included. However, these terms were dropped in the stepwise procedure for the logit model and replaced by the terms which are elaborated upon below.

The total net worth of the firm (TNW) gives a good indication of the value of the firm and the type of debt it can handle. Total net worth is defined as the total of the company stock (common and preferred) and its retained earnings. This is presented in the balance sheet as the total under Shareholder's Equity. This value is a measure of the financial strength of the firm. The significance of this variable is not surprising given that, in the extension of credit, the stronger the firm, the less the risk.

The interaction of the branch of bank (BRN) and the total net worth indicates the importance of TNW in the decision process; it further implies that this importance varies among branches. One reason could be that the branch managers place varying emphasis on the different levels of total net worth. A second possible conclusion is that the size of the firms vary in different regions for the different commercial banking centres, resulting in a high disparity of total net worth across the branches. This would then further result in a significant interaction term. We should also take note of the many hidden factors, already discussed, associated with the BRN variable. For example, it is possible that the type of industry may interact with the TNW values. This factor may play a leading role in the interaction of BRN and TNW.

The last term, amount of loan divided by total net worth (AMT/TNW), is an extremely important variable. This result seems logical due to the fact that the amount of loan and the total net worth usually vary directly. If the amount of the loan is large, one would expect the total net worth to be relatively large in order to indicate that the firm is financially secure and capable of supporting a large debt. Conversely, if the amount of the loan is relatively small, then the total net worth does not have to be as high. In other words, for a successful applicant, one would expect to see an upper bound to the size of the ratio of AMT/TNW, regardless of the size of the loan. The value of this ratio represents the financial strength of the firm, and its ability to handle the accompanying debt - the smaller the ratio, the greater the financial strength.

The value of the coefficient of this variable (AMT/TNW) is positive (0.001377). Due to the relationship between this variable and the probability of a loan (inversely related), one would expect the coefficient to be negative. This appears curious, however, it could be that an explanation is revealed from the analysis of the raw data. The residual analysis (in Section 5.2; Table A-6.1) illustrates that there exists one outlier where the ratio of the AMT/TNW value is very high yet the applicant was still granted the loan. There were also four cases where the applicant firm had negative TNW and the loan was rejected. (Remember, the lower the TNW, the lower the ratio.) The negative value has the effect of pushing the ratio to a small number. This results in cases where the ratio is small yet the application is denied. As a result, the outlier in conjunction with these four other cases where the firm had negative TNW (and the loan application was rejected) had the effect of changing the slope of this interaction term from negative to a significant positive value. Although, the coefficient of this interaction term is affected, this does not detract from the methodology which is employed for the analysis. We discuss the problem of the outlier more fully in the section on residual analysis.

The models in the literature appear to incorporate variables similar to those used in this thesis. As a comparison, the Doreen and Farhoomand (1983) study of commercial loan applications produced the following significant variables:

(i) competitive situation;
(ii) cash flow projection;
(iii) owner/manager experience;
(iv) equity;
(v) overall security;
(vi) product (industry); and
(vii) communication skills.

We also refer to the Altman study (1968), which was presented in Section 2.6.

All these models consider equity or the total net worth (TNW). The coefficients for these studies are positive. In other words, the higher the equity, the greater the likelihood of a loan approval. Also, the models take into consideration the industry or competitiveness of the business (level of BRN and NEW variables). Further comparisons can be examined; however, it is difficult to conduct this examination in any great detail due to the different nature of the studies.

5.1.4 The Logistic Regression with Mixed Effects Model

As was outlined in Chapter Four, the predictor variables included in the "best" model for the logit regression with fixed effects, described in Section 5.1.3, were used for the logit regression model with mixed effects. The reason for this procedure is that an extensive and cost efficient method for variable selection for the random effects is not available at this time. Consequently, the model which is used is the following:

$$\text{logit}(p) = b_0 + BRN_i + TY_j + NEW_k + BRN_i * TNW + b_1 * AMT + b_2 * AMT/TNW \quad (5.1.4)$$

where $BRN_i \sim N(0, \sigma_1^2)$,
$BRN_i * TNW \sim N(0, \sigma_2^2)$; and

the remaining variables are as defined in the fixed effects model. Thus, the branch effect is considered random. (The reasons are discussed in detail in Section 4.1.)

One main concern in this random effects model is the variability of the branch effect. The variance is estimated to be 1.89, using the EM algorithm (discussed in Chapter 4). In order to assess the size of this variance, we can create an interval for the branch effect. Since it is easier to relate to a probability, this interval can then be translated into an interval for the p's (when all the other variables in the model are at their mean levels). Consequently, an interval covering plus two and minus two standard deviations of the branch effect ($2 \times \sqrt{1.89}$) for p is as follows:

(.2467, 0.9877)

where the mean value for p is 0.8366. This interval would cover approximately 95 percent of the probabilities when other conditions are constant. Can we infer from this that there exists a large variability from branch to branch in the decision process at the bank? It appears that the variance is very large causing p to fluctuate over almost the entire range of possible values. This is attributed to the sampling scheme as opposed to the actual variability of the decision process across the branches. In Section 3.5, we stated that the sampling procedure imposed on the research would result in the significance of this branch effect and would, consequently, cause the variance to be large. However, this fact does not detract from the rest of the analysis. All the remaining variables are viable and we do well to continue the model building. This problem is addressed in more detail in Section 5.8. The parameter estimates as well as the coefficients for this model are given in Table A-4 under the heading LRME.

As discussed in Section 5.1.3, in 31 of the branches, all the observations are either yes or no - but not both. This resulted in an estimation problem for the fixed effects model. However, for the model where the branch effects are considered random, this problem does not exist. The branch effect estimates are not forced to infinity, and the model handles the data well. These estimates do not rely solely on the observations but also on the information shared across the cells. Consequently, each branch effect estimate, regardless of the actual number of observations for each branch, is a result of information concerning the yes's and the no's.

In addition, the variance of the interaction term (the branch effect interacting with the total net worth of the firm) was calculated. This value is 0.0000042. Again, we can interpret this value. Consider all the variables in the model to be at their means; we select BRN 1 for illustration. Let this interaction effect vary by plus two standard deviations and increase the value of TNW by $10,000. This, then, has the effect of increasing the odds ratio [p/(1-p)] by a factor of 0.413 percent. On the other hand, if we vary this interaction effect by minus two standard deviations, the odds ratio is decreased by 0.407 percent. Thus, the fluctuation of this interaction effect has very little visible effect on the outcome of the odds ratio of p.

5.2 Results of the Residual Analysis

In order to evaluate, in detail, the fit of the 3 models (LDA, LRFE and LRME) to the data, residual analysis was performed. Residual analysis for logit regression is not as straightforward as that for linear regression. In linear regression, the value of the outcome variable is continuous. The predicted value provided by the model is compared to the observed value for the given data points. The difference between the observed and the predicted is the residual value, which is subject to specific interpretations. Most analyses involve an investigation of the existence of visible patterns exhibited by these residuals (Neter, Wasserman and Kutner, 1985). In logistic regression, the outcome variable is [0,1]. The predicted values are probabilities, and, thus, the value of the residual ($r_i = y_i - \hat{y}_i$) is limited in range. Consequently, inferences from this type of analysis are difficult to perform.

An alternative approach is presented by Landwehr, Pregibon and Shoemaker (1984). In their paper, the authors propose three graphical methods that can be used to assess the fit of logistic regression models. The authors refer to "local mean deviance plots," which were employed to detect overall lack of fit. Then, "empirical probability plots" were used to point out isolated departures from the fitted models. Finally, residual plots (referred to as "partial residual plots") were used to identify specific causes of lack of fit. These techniques were applied to a data set concerning the survival of patients who had undergone surgery for breast cancer. The results from this analysis show the ability of these three graphical methods to detect lack of fit and, consequently, build better models.

One of the drawbacks of using this method is that it is extremely costly. The calculations needed to create the graphical results require a significant amount of computer time. Second, in order to expose any visible trends, the "local mean deviance method" requires a clustering algorithm, and the "partial residual plots" require the specification of a smoothing algorithm (Rubin, 1984). Since the selection of these algorithms is arbitrary - there are no preferred ones - the results of the Landwehr et al study (1984) may be dependent on the particular algorithms chosen. For these reasons, an alternative method for residual analysis has been selected.

We follow a method proposed by Rubin (1984). In order to clearly illustrate this procedure, we briefly introduce the "Theorem on Propensity Scores", developed by Rosenbaum and Rubin (1983) for observational studies. Let us consider a study where there are two outcome groups: (a) the loan application is approved (yes group - y=1) and (b) the loan application is rejected (no group - y=0). Also, consider a vector of independent variables represented by **x**. The theorem states that **x** and y are conditionally independent given the propensity score (p_x), where the propensity score is defined as

$$p_x = \text{prob}\ (y=1 \mid x)\ ,$$

in other words, the posterior probabilities. That is, the probability distribution function (p.d.f.) of **x** given p_x and (y=1) should be the same as the p.d.f. of **x** given the propensity score and (y=0). (See Rosenbaum and Rubin, 1983, for a full presentation of this theorem.)

Given this result, the residual analysis is as follows: If $\hat{p}_x$ is a good estimate of the probability of (y=1) given **x** [prob(y=1 | x)], then the regression of **x** on $\hat{p}_x$ should be essentially the same for the y=1 and y=0 groups (Rubin, 1984). One way to illustrate this relationship is to plot each independent variable in the model(s) against the estimated probabilities, $\hat{p}$. If the plots for the yes group (y=1) are approximately the same as for the no group (y=0) across all the variables, then the model should provide a good fit to the data.

It should be noted that one of the benefits of using this "propensity score" method for residual analysis is that it allows for a direct comparison with the logit and discriminant models; this benefit is not readily available in other types of methods.

The propensity score estimate ($\hat{p}_x$) is easily ascertained from the logit model. In this model, we are predicting the value of p. In the LDA model, on the other hand, the calculation of p is not as obvious. In this model, p is the posterior probability calculated as a by-product of the discriminant analysis. In other words, after the analysis has been completed, p represents the posterior probability of the approval of a particular case. The posterior probability is obtained as a function of the discriminant score in the following manner:

$$p = 1/(1 + q\exp\{-Z_{ijkl} + t^*\})$$

where p is the posterior probability of a particular case ijkl;

q is the ratio of prior probabilities of group association;

Z_{ijkl} is the discriminant score of this case; and

t^* is the cutoff value determined to discriminate between the two groups (Green, 1978).

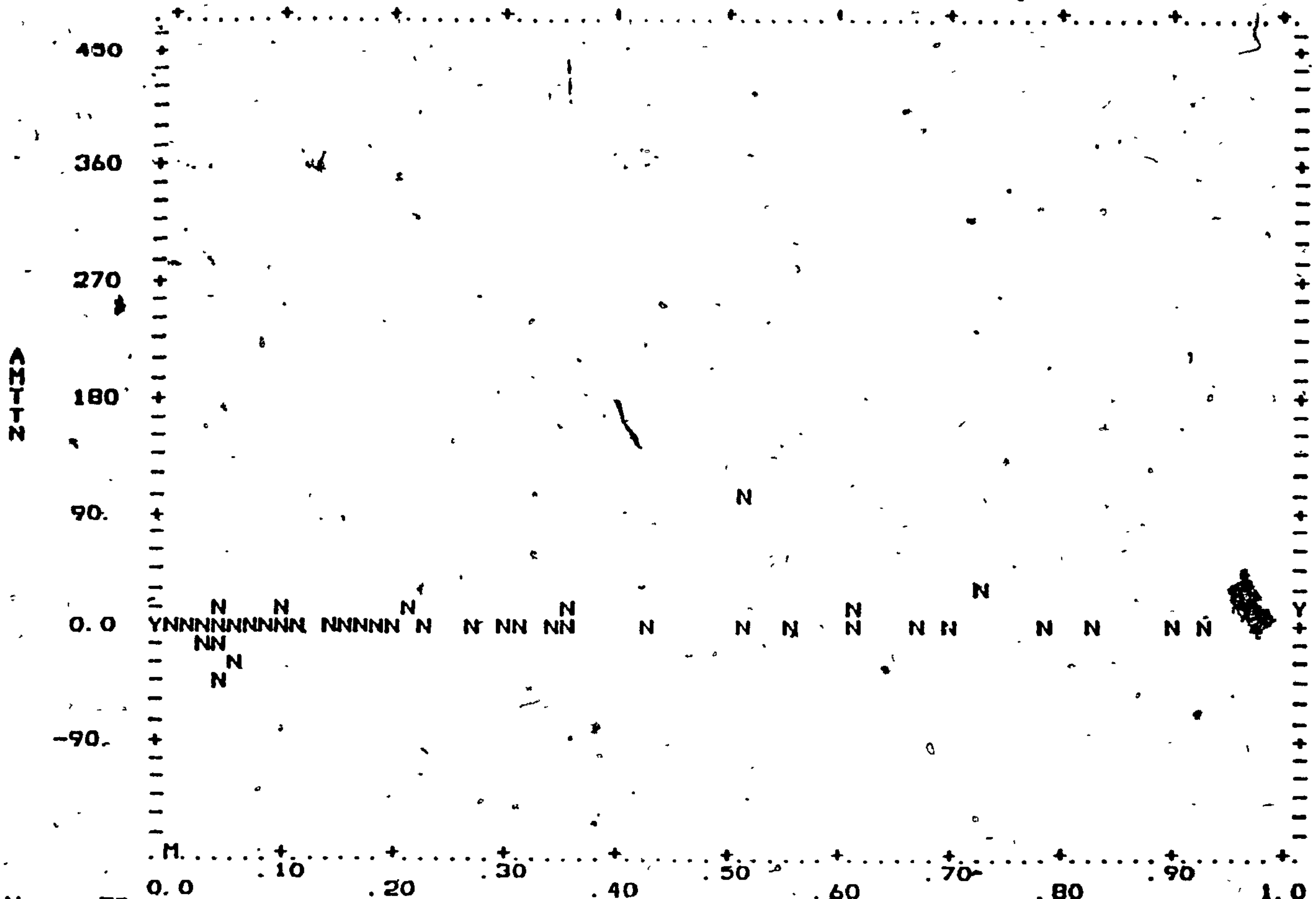
In the analysis, the continuous variables in the models were first plotted versus the probability level and the outcome. These results are shown in Table A-6.1. Here, the no group and the yes group are plotted both separately and together. (The results for LRFE and LRME are given in Tables A-6.1.1 and A-6.1.2 respectively. As well, for illustration, the tables for AMT/TNW are reproduced below. The only continuous variable in the LDA model is the AMT variable interacting with the BRN effect. There was no residual analysis performed on the BRN*AMT and BRN*TNW interaction terms, for any of the models, due to the extremely large number of categories which made analysis impractical. Thus, the remaining categorical variables for LDA are examined in Tables A-6.3 and A-6.4 only.) Recall that, according to the theorem given above, the scatter plots for the yes and no groups should be similar for a given level of p. But, an examination of the scatter plots reveals little conclusive evidence. Most of the observations for the no group have small p values (probabilities) and for the yes group the probabilities are high (near one). Given the small data set, it is not easy to compare the different distributions for the various intervals of p. In conclusion, there are no obvious departures from "similarity" that would call into question the appropriateness of the models.

One point, which should be mentioned, is the outlier observed in the yes group for the AMT/TNW (or AMTTN) term in the logit models. Recall that, in Section 5.1.3, we found the coefficient of this term to be positive (logically, it should be negative). The presence of the outlier along with the four cases where the TNW value was negative had an effect on this value. To investigate this further, the logit model with fixed effects (LRFE) was run again on 282 cases (the outlier was excluded). The AMT/TNW term was no longer significant; however, the coefficient was still positive. This implies that the cases where TNW was negative had an observable effect on the model. However, the classification results were not significantly different from the full model on the full data set. (See classification results in Table A-7 and Section 5.3.)

Frequency histograms were then created for all the continuous variables in the LRFE and LRME models, which were categorized both by the outcome level (no or yes), and by the level of probability (divided into five categories: 0-.2, .2-.4, ...). The results are given in Tables A-6.2.1 (LRFE) and A-6.2.2 (LRME). The tables for AMT are reproduced here for illustration.

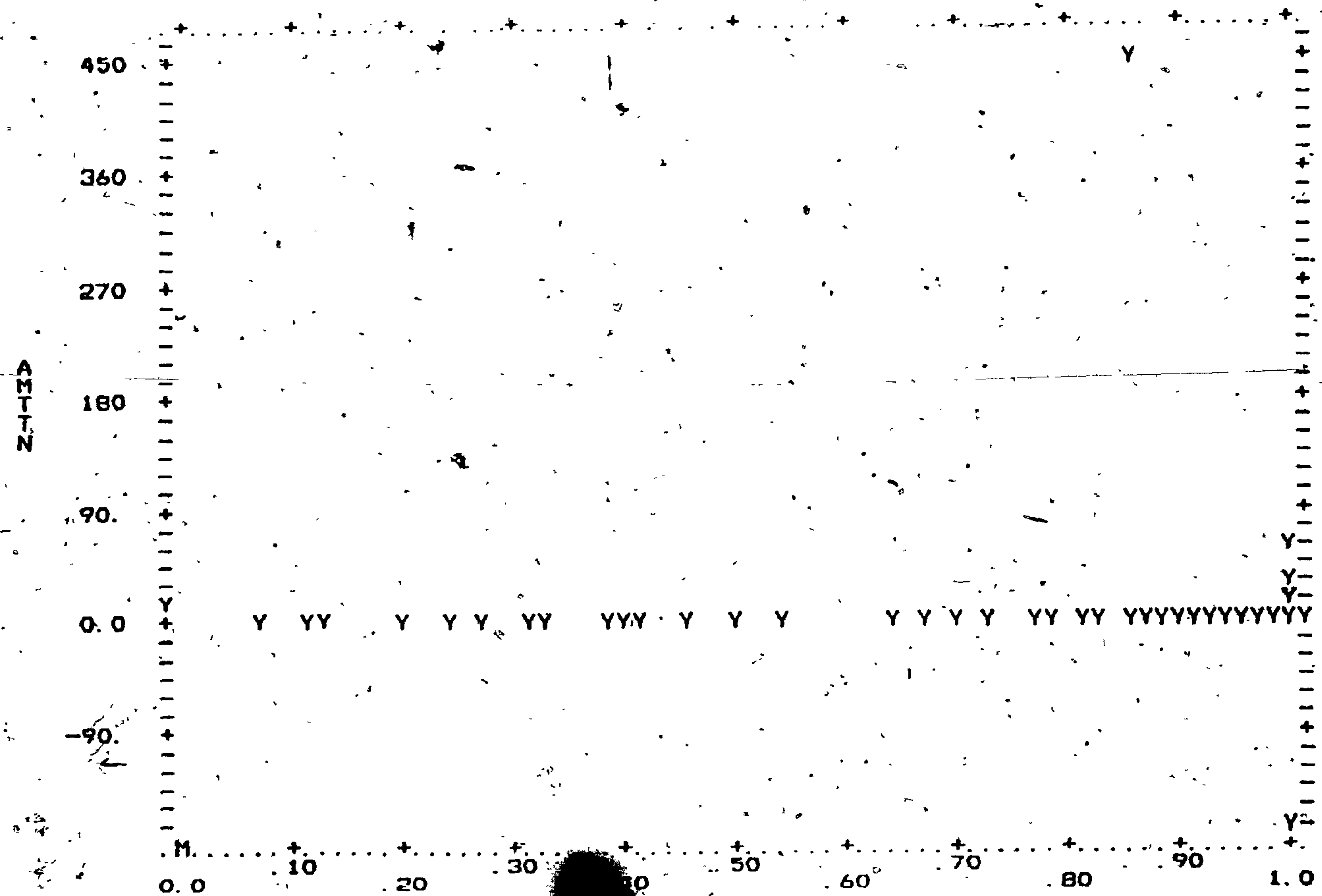
Table A-6.1.1

AMT/TNW for LRFE



N = 79
R = .2353
P = .0366

```
--REGRESSION LINE--     --RMS--          MEAN      S.D.
Y=-1.3100 +13.081*X      208.44      X   .23643    .26551
                                     Y  1.7027    14.759

LOGP   .  VERSUS AMTTN      ( 28 VS. 31 )   GROUP=NO          SYMBOL=N
1PAGE  25  BMDP6D PLOTS ANAL OF BDATA
```

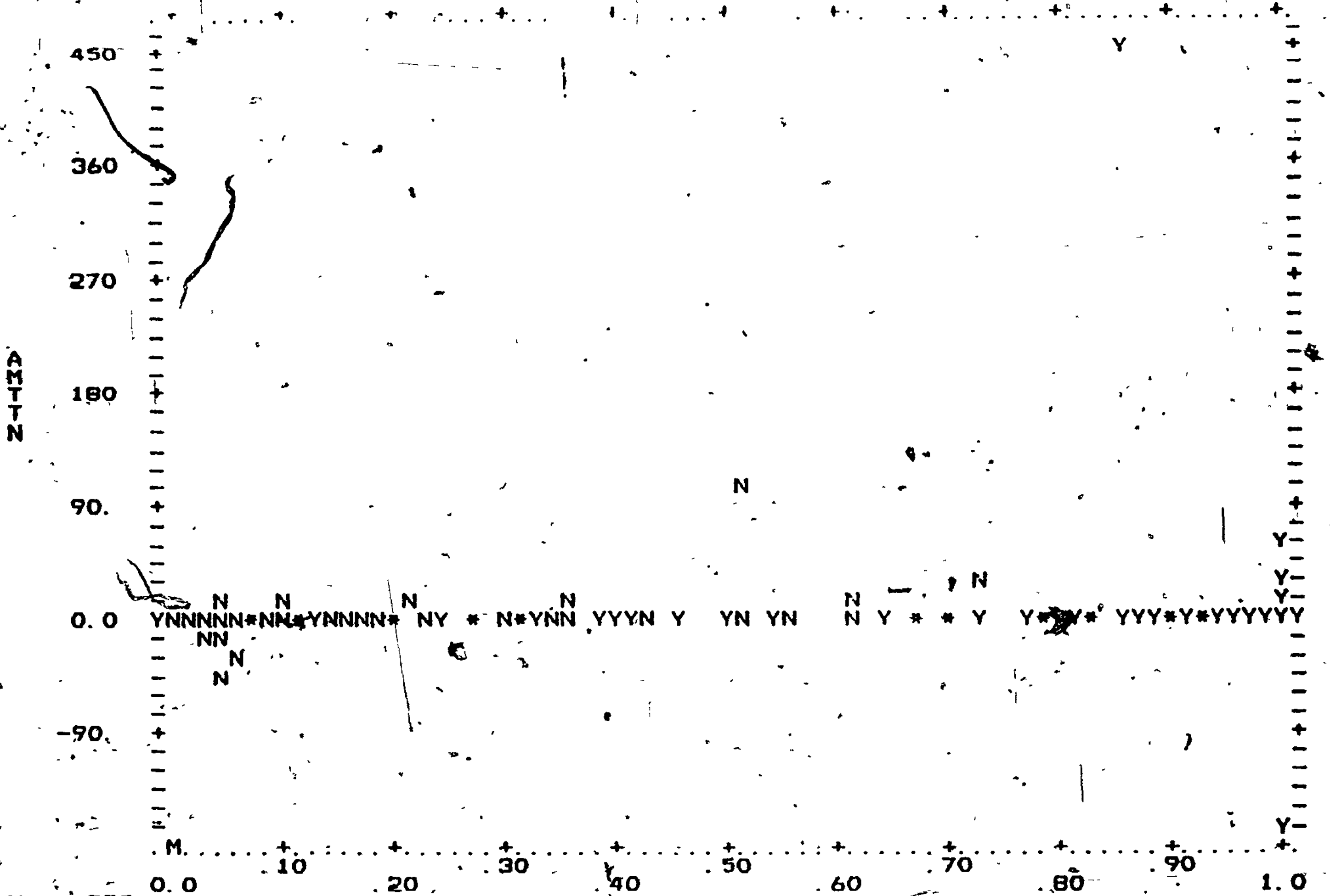


N = 204
R = -.0254

P = .7189

			MEAN	S.D.
--REGRESSION LINE--	--RMS--	X	.90861	.19148
Y= 7.9231 -4.4983*X	1156.4	Y	3.8359	33.933

LOOP VERSUS AMTTN (28 VS. 31) GROUP=YES , SYMBOL=Y
1PAGE 26 BMDP6D PLOTS ANAL OF BDATA



N = 283
R = .0376
P = .5294

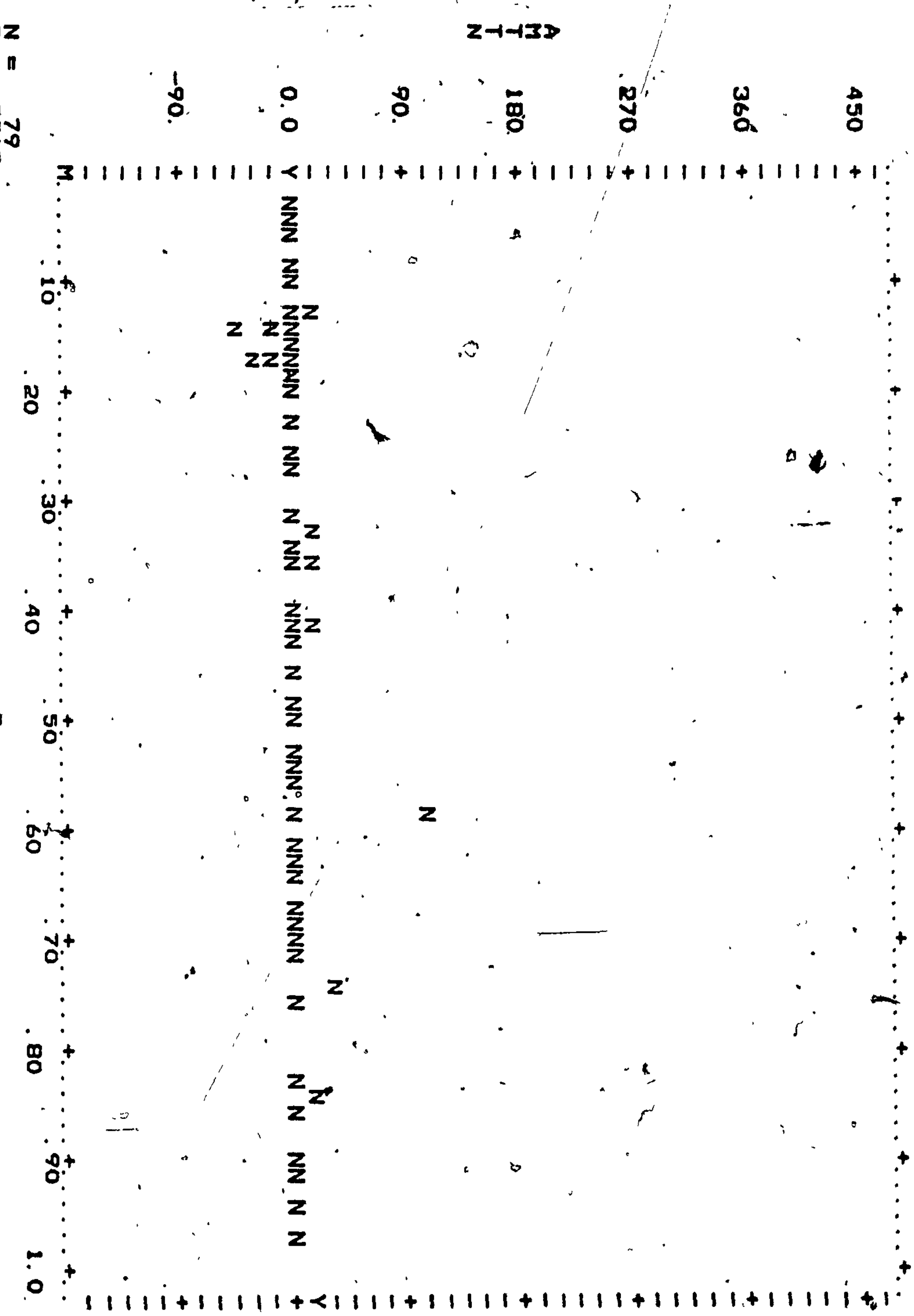
--REGRESSION LINE--	--RMS--		MEAN	S. D.
Y= 1.0808 +3.0264*X	891.91	X	.72097	.37032
		Y	3.2627	29.833

LOGP VERSUS AMTTN (28 VS. 31) GROUP=NO , SYMBOL=N
GROUP=YES , SYMBOL=Y

1PAGE 27 BMDP6D PLOTS ANAL OF BDATA

Table A-6.1.2

AMT/TNW for LRME

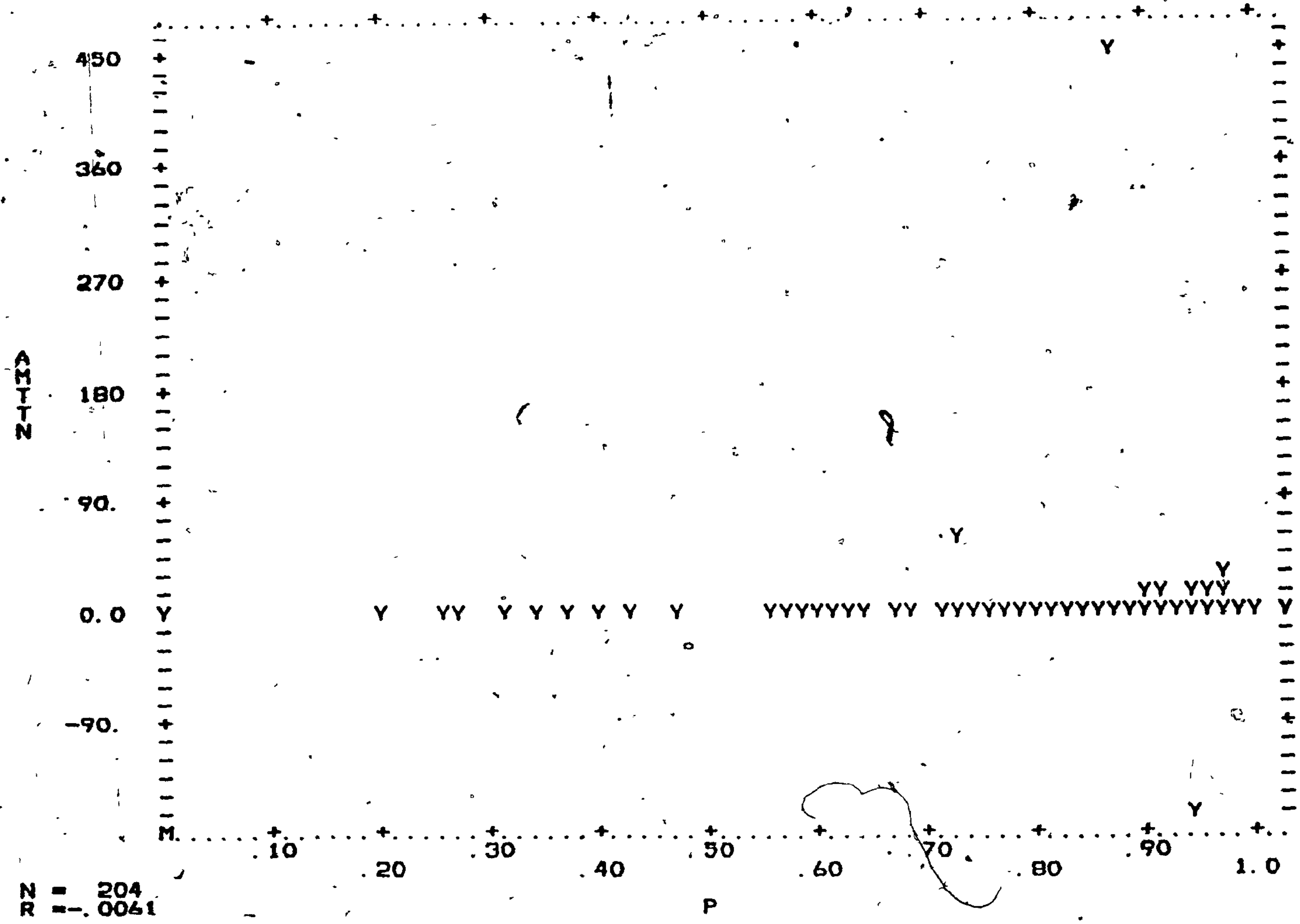


N = 79
R = .2365
P = .0356

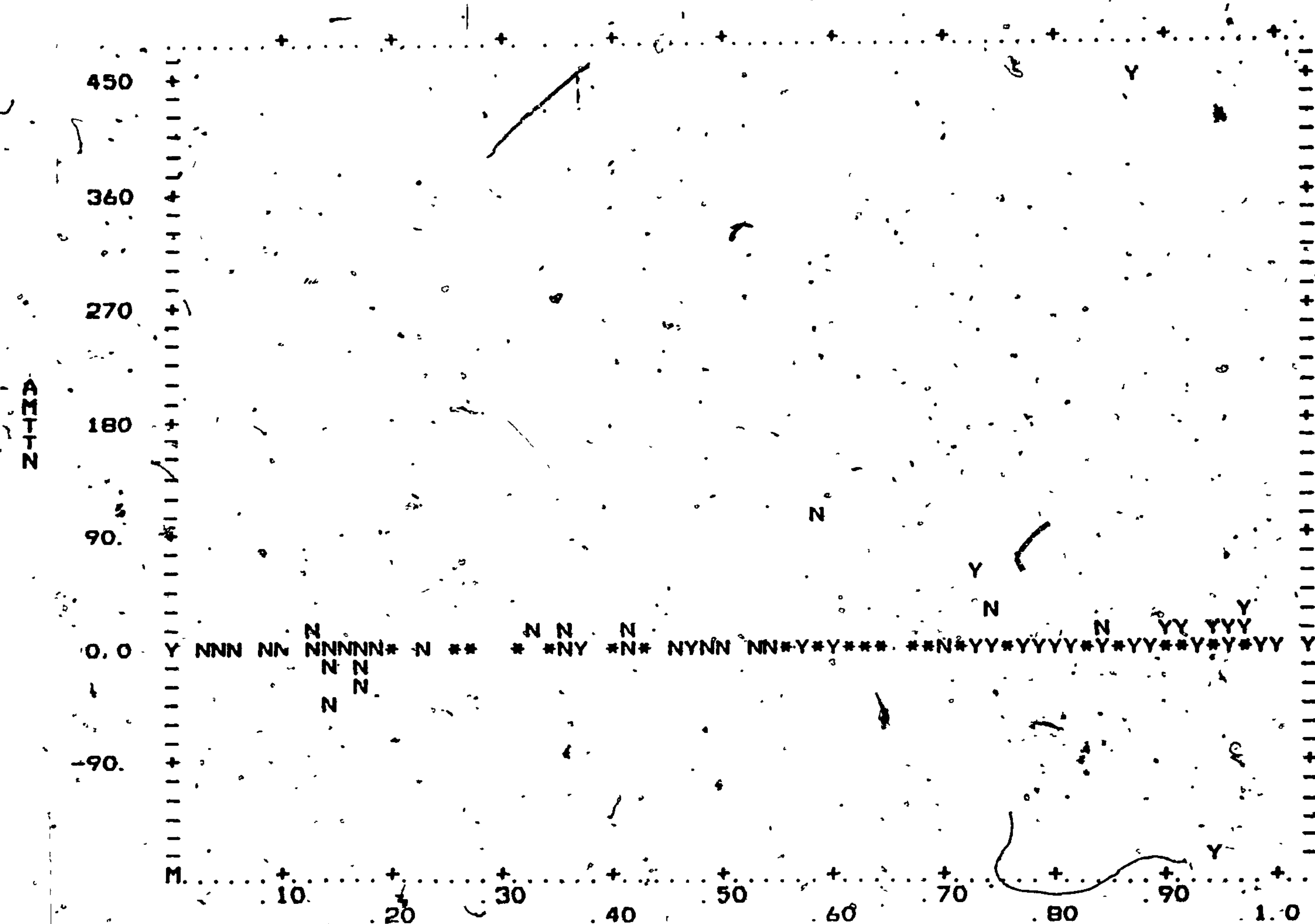
--REGRESSION LINE--	--RMS--
Y=-3.4971 +13.638*X	208.32

	MEAN	S.D.
X	.38715	.25591
Y	1.7827	14.759

LRPP VERSUS AMTTN (29, VS. 31) GROUP=NO , SYMBOL=N
1PAGE 25 BMDP6D PLOTS ANAL OF BDATA




```
N =  204
R =-.0061

P = .9305

                                                MEAN      S.D.
--REGRESSION LINE--     --RMS--       X    .84994    .16336
Y= 4.9216 -1.2774*X      1157.1       Y    3.8359    33.933

LRPP      VERSUS AMTTN        ( 29 VS. 31 )   GROUP=YES        SYMBOL=Y
1PAGE  26  BMDP6D PLOTS ANAL OF BDATA
```



N = 283

R = .0489
P = .4125

P

--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y=-.44394 +5.1428*X	891.03	X	.72075	.28386
		Y	3.2627	29.833

LRPP VERSUS AMTTN (29 VS. 31) GROUP=NO , SYMBOL=N
GROUP=YES , SYMBOL=Y

1PAGE 27 BMDP6D PLOTS ANAL OF BDATA

The histograms for TNW and AMT/TNW (AMTTN) are fairly straightforward. The distributions for the no and yes groups are similar for both the LRFE and LRME models. Let us examine the amount of loan variable (AMT). For both the fixed effects and mixed effects models, and for both the yes and the no groups, the mean value of the amount gets larger as the value of p (the probability that a loan application is accepted) gets larger. As well, if we consider the mean values for the two groups within each category of p, we see that the values are very close together for almost all of the categories. As a result, based on the propensity theorem, we can conclude that the distributions for the no and yes groups are similar given the level or category of p. This implies that the models provide a good fit to the data.

Third, three-way frequency tables were charted for each independent variable, using the probability level (as described above) and the outcome (no or yes group). The continuous variables were made discrete for this analysis. The full log-linear models were fit for each table, and the three-factor interaction and then one two-factor interaction were tested for significance. The major emphasis of this residual analysis is on this two-factor interaction term. The propensity theorem states that the distribution of the independent variable is conditionally independent of p_x (the propensity score) given the value of y (the group association - yes or no group). Thus, the two-factor interaction between the independent variable (e.g., amount of loan - AMT) and the value of y should be tested for significance in the hierarchical model. This was done for each independent variable in all three models. If this two-factor interaction term is not significant, this implies that the distributions of the independent variables, given p, are the same for the y = 1 and y = 0 groups. In other words, the models presented in Section 5.1 are appropriate.

The preliminary step in this analysis is to test for the three-factor interaction. In a hierarchical model such as this one, if this term is present, all the two-factor interactions will be present. The initial results indicate a consistent significance of this three-factor interaction term. The results for the three models, LDA, LRFE and LRME, are as follows:

Models	Variables	p-value
1.LDA	TY	.00593
	NEW	.00071
	BRN	.00778
	TYNEW	.00910
2.LRFE	TY	.46558
	NEW	.00161
	BRN	.00067
	AMT	.12685
	AMTTN	.00000
3.LRME	TY	.00000
	NEW	.64149
	BRN	.52399
	AMT	.33484
	AMTTN	.00000

More detailed results are shown in Table A-6.3 in the appendix.

Table A-6.2.1

AMT for LRFE

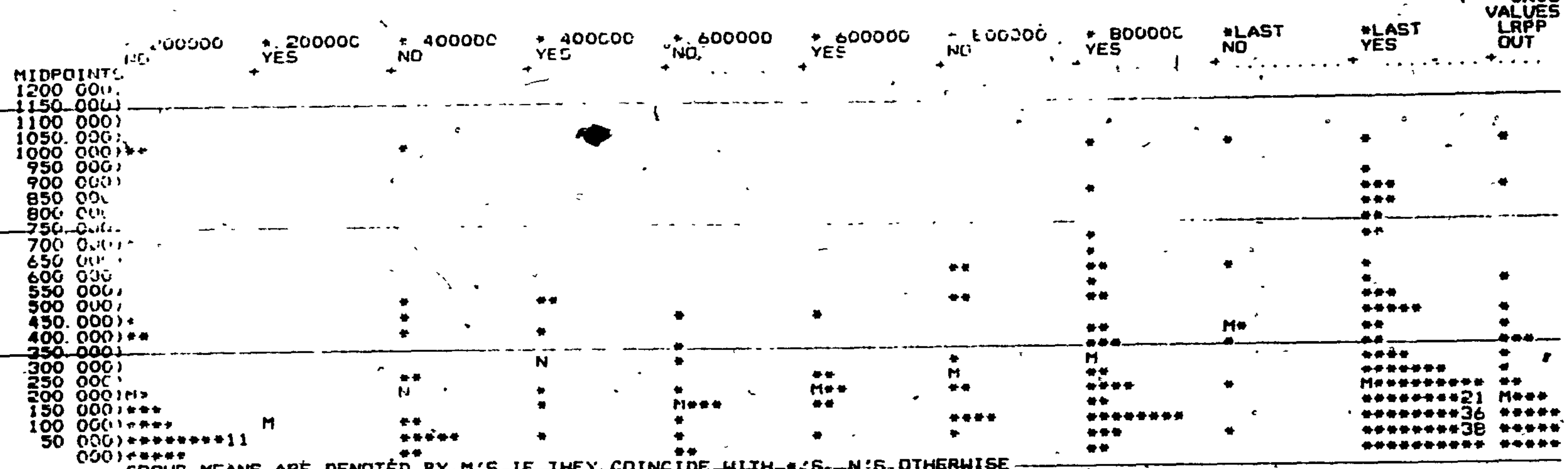
```
PAGE   4  BMDP7D PLOTS ANAL OF BDATA

             *************                            *************
HISTOGRAM OF * AMT        *  (    3)   GROUPED   BY   * LOOP      *  (   28)
             *************                       AND  * OUT       *  (    2)
                                                      *************

          * 200000    * 200000    * 400000    * 400000    * 600000    * 600000    * 800000    * 800000    *LAST       *LAST
          NO          YES         NO          YES         NO          YES         NO          YES         NO          YES
MIDPOINTS +           +           +           +           +           +           +           +           +           +           +
 1200.000)
 1150.000)
 1100.000)
 1050.000)
 1000.000)***         *                       *                                                           *           *
  950.000)
  900.000)                                                                                                            *
  850.000)*                                                           *                                               ***
  800.000)                                                                                    *                       **
  750.000)                                                                                                            **
  700.000)*                                                                                                           ***
  650.000)*                                                                                                           *
  600.000)                                                                        ***                                 ***
  550.000)            *                                                                                   N           **
  500.000)*           *                       *           *                       *           *                       ****
  450.000)            *           **                                              +           *                       *****
  400.000)****        **                                                          N                       *           ****
  350.000)**          **          *                                                                                   *****
  300.000)            *                       N           *           M           *                                   ****
  250.000)*           *           *                                               *           N                       *********11
  200.000)M**         **          N           **          M           *           *           ****        *           M********11
  150.000)****        M**         **          **          *                       *           **                      *********22
  100.000)*********   *******                             ***         *                       *                       *********42
   50.000)*********18 *********18 ***         **                      *                       *                       *********39
     .000)*********11 *********11 *                                                                                   *********12
          GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE

MEAN         189.966     136.843     195.100     282.500     198.857     289.600     398.444     267.000     528.333     203.328
STD DEV      265.872     186.590     171.003     322.988     155.806     329.497     181.480     224.243     403.859     216.747
S  E  M       34.911      26.128      54.076     114.194      58.889     147.356      60.493      67.612     233.168      16.292
MAXIMUM     1000.000    1000.000     450.000    1000.000     500.000     850.000     602.000     800.000     980.000    1000.000
MINIMUM        7.000       5.000      20.000      35.000      78.000      35.000     150.000      25.000     202.000       8.000
CASES INCL        58          51          10           8           7           5           9          11           3         177

                    ALL GROUPS COMBINED
               (EXCEPT CASES WITH UNUSED VALUES
             FOR VARIABLES LOOP     AND OUT     )
                  MEAN                203.968
                  STD  DEV            229.046
                  S  E  M              12.440
                  MAXIMUM             10.0E+2
                  MINIMUM               5.000
                  CASES EXCLUDED     (      0)
                  CASES INCLUDED          339

                  ROBUST S.D          215.047
```

Table A-6.2.2

AMT for LRME

PAGE UUDF7! PLOTS ANAL OF BDATA

```
HISTOGRAM OF  ************        GROUPED  BY   ************
              * AMT      *  (  3)           AND  * LRPP     *  ( 29)
              ************                       * OUT      *  (  2)
                                                 ************
```



GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE

	.000000 NO	.200000 YES	.400000 NO	.400000 YES	.600000 NO	.600000 YES	.800000 NO	.800000 YES	*LAST NO	*LAST YES	CASES UNUS VALUES LRPP OUT
MEAN	196 156	120 000	221 667	297 500	169 500	201.667	268 923	295 361	419.571	192.770	132
STD DEV	277 779	0 000	273.834	194 004	137 370	113 285	207.737	253.085	294.020	219.672	198
S E M	49 105	0 000.	70 704	79 202	39 655	37 762	57 616	42. 181	111. 129	17 818	26
MAXIMUM	1000 000	120 000	1000 000	500 000	450 000	450 000	602 000	1000 000	980. 000	1000 000	1000
MINIMUM	10 000	120 000	11 000	50 000	15 000	35.000	25 000	8. 000	70. 000	10 000	5
CASES INCL	32	1	15	6	12	6	13	36	7	152	

ALL GROUPS COMBINED
(EXCEPT CASES WITH UNUSED VALUES
FOR VARIABLES LRPF AND OUT

MEAN	218 162
STD DEV	232 388
S E M	13 814
MAXIMUM	10 0E+2
MINIMUM	5 000
CASES EXCLUDED	(0)
CASES INCLUDED	283

ROBUST S-D - - 266 44

It should be mentioned that for all the variables, there were many cells with small expected values (< 1) and many cells equal to zero. Although it is not clear how many small cells are permitted, many authors (see, for example, Conover, 1980) state that no more than twenty percent of the cells should have expected values less than five. Otherwise, the results of a chi-square test become questionable. A more detailed analysis would include an examination of the two-factor interaction term in order to attain less ambiguous results.

As can be seen from the results reported above, many of the two-factor interactions will be present due to the significance of the three-factor interactions. In order to understand this result better, residual analysis was performed on the log-linear models. This is shown in Table A-6.4. (Table A-6.4.1 for LDA; Table A-6.4.2 for LRFE; and Table A-6.4.3 for LRME.) The tables of standardized deviates are largely comprised of mild departures (less than plus and minus two), which implies a reasonable fit is provided by the models. (These tables are reproduced below.) We examine the two-factor interaction in more detail here:

Examination of Two-Factor Interaction

Models	Variables	p-value
1.LDA	TY	.0164
	NEW	.0001
	BRN	.1526
	TYNEW	.0178

2.LRFE	TY	.5282
	NEW	.0003
	BRN	.2148
	AMT	.0990
	AMTTN	.0348
3.LRME	TY	.0000
	NEW	.8456
	BRN	.9948
	AMT	.3448
	AMTTN	.0008

The LDA Model

The variable TY had a significant TY-OUT (type of loan - outcome) two-factor interaction term. An examination of the tables of standardized deviates (Table A-6.4) indicates that there are eight cells with absolute values greater than 1.0 (the largest being 1.8). The observed values in these cells totalled 26. Thus, 26 out of 339 cases are responsible for the low p-value as well as the significance of this interaction term.

TWO-FACTOR INTERACTION - LDA MODEL

PAGE 13 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.1	0.0	0.2	-0.9
	.400000	0.8	-0.3	0.0	-0.4
	.600000	0.2	-1.3	1.4	-0.8
	.800000	0.0	0.3	0.0	-0.6
	LAST	-0.6	1.7	-1.7	0.5
YES	.200000	-0.2	0.0	-0.4	1.8
	.400000	-1.2	0.4	0.0	0.5
	.600000	-0.2	1.2	-1.3	0.7
	.800000	0.0	-0.3	0.0	0.5
	LAST	0.2	-0.5	0.5	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.1	0.0	0.4	-1.1
	.400000	0.8	-0.1	0.0	-0.3
	.600000	0.3	-1.5	1.3	-0.7
	.800000	0.0	0.4	0.0	-0.6
	LAST	-0.5	1.5	-2.6	0.6
YES	.200000	-0.1	0.0	-0.3	1.1
	.400000	-1.5	0.5	0.0	0.6
	.600000	-0.1	1.1	-1.3	0.7
	.800000	0.0	-0.1	0.0	0.6
	LAST	0.2	-0.5	0.5	-0.1

PAGE 17 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.3	0.5	-1.3	0.0
	.400000	0.7	-0.6	0.0	0.6
	.600000	1.2	-1.0	-0.8	-0.6
	.800000	-0.6	-0.2	0.0	0.9
	LAST	-1.5	2.3	-0.4	-1.2
YES	.200000	2.6	-1.1	2.6	0.0
	.400000	-1.0	0.8	0.0	-0.8
	.600000	-1.1	0.9	0.7	0.5
	.800000	0.5	0.1	0.0	-0.8
	LAST	0.4	-0.6	0.1	0.3

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.7	0.6	-1.7	0.0
	.400000	0.7	-0.5	0.0	0.6
	.600000	1.2	-1.0	-0.7	-0.5
	.800000	-0.6	0.0	0.0	0.8
	LAST	-1.6	2.0	-0.2	-1.3
YES	.200000	1.6	-1.1	1.6	0.0
	.400000	-1.2	0.8	0.0	-0.9
	.600000	-1.1	0.9	0.7	0.6
	.800000	0.6	0.3	0.0	-0.8
	LAST	0.4	-0.6	0.2	0.4

STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	0.0	0.2	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.4	0.0	0.0	0.0	-0.5	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.2	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0	0.0
	LAST	0.8	1.0	0.0	-0.4	0.2	-0.6	-0.5	-0.5	-0.4	-0.7	-0.6	-0.7
YES	.200000	0.0	0.0	-0.4	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	-0.6	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	-0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0
	LAST	-0.2	-0.3	0.0	0.1	-0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.2

OUT	LDAP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.7
	.800000	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.0	-0.2	-0.3	-0.9	-0.6	-0.3	-0.3	0.8	-0.9	-0.4	1.2	3.3
YES	.200000	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	-1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	0.0	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.0	0.0	0.1	0.3	0.2	0.1	0.1	-0.2	0.3	0.1	-0.4	-1.0

OUT	LDAP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.7
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0
	LAST	-0.6	0.5	0.8	0.0	-0.3	-0.3	-0.6	1.2	-0.7	-0.5	-0.4	0.8
YES	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	-0.7	0.0	0.0	0.0	0.0	0.0	[illegible]	0.0	0.6
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
	LAST	0.2	-0.1	-0.2	0.0	0.1	0.1	0.2	-0.4	0.2	0.1	0.1	-0.2

OUT	LDAP	BRN 47.0000	48.0000
NO	.200000	0.0	0.0
	.400000	-0.8	0.0
	.600000	0.0	0.0
	.800000	0.2	0.0
	LAST	3.3	-0.5
YES	.200000	0.0	0.0
	.400000	1.2	0.0
	.600000	0.0	0.0
	.800000	-0.2	0.0
	LAST	-1.0	0.1

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	TYNEW 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	11.0000
NO	.200000	0.0	-1.3	0.0	0.0	0.0	0.5	0.0	0.2	0.0	-0.9	0.0	0.0
	.400000	0.0	0.6	0.4	0.0	0.0	0.0	-0.8	0.0	-0.4	0.0	0.0	0.0
	.600000	-0.7	0.8	0.1	1.2	-0.3	0.8	-1.3	0.0	-0.7	-0.6	-0.7	0.0
	.800000	0.0	0.0	-0.6	0.0	0.0	0.0	0.2	0.0	-0.6	0.0	0.0	0.0
	LAST	-0.9	-1.0	0.4	-1.3	-0.8	0.9	2.2	-0.9	1.8	-0.5	-0.5	-0.7
YES	.200000	0.0	2.6	0.0	0.0	0.0	-1.1	0.0	-0.4	0.0	1.8	0.0	0.0
	.400000	0.0	-0.8	-0.6	0.0	0.0	0.0	1.2	0.0	0.5	0.0	0.0	0.0
	.600000	0.6	-0.7	-0.1	-1.1	0.2	-0.7	1.2	0.0	0.6	0.5	0.6	0.0
	.800000	0.0	0.0	0.5	0.0	0.0	0.0	-0.2	0.0	0.5	0.0	0.0	0.0
	LAST	0.3	0.3	-0.1	0.4	0.2	-0.3	-0.6	0.2	-0.5	0.1	0.1	0.2

OUT	LDAP	TYNEW 12.0000	13.0000	14.0000	15.0000
NO	.200000	-0.9	0.0	0.0	0.0
	.400000	0.0	0.6	0.0	0.0
	.600000	0.0	-0.6	-0.7	0.8
	.800000	0.0	0.0	0.9	0.0
	LAST	0.7	-1.3	0.4	-0.3
YES	.200000	1.8	-0.0	0.0	0.0
	.400000	0.0	-0.8	0.0	0.0
	.600000	0.0	0.5	0.6	-0.7
	.800000	0.0	0.0	-0.8	0.0
	LAST	-0.2	0.4	-0.1	0.1

TWO-FACTOR INTERACTION - LOGIT FIXED EFFECTS

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOGP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.8	-1.2	-0.1	-0.3
	.400000	0.4	0.6	0.6	-1.0
	.600000	-0.2	0.2	0.0	0.0
	.800000	-0.1	-0.3	-0.7	0.5
	LAST	-0.9	0.7	-0.8	0.8
YES	.200000	-0.8	1.3	0.1	0.4
	.400000	-0.5	-0.7	-0.7	1.1
	.600000	0.3	-0.2	-0.1	0.0
	.800000	0.1	0.3	0.6	-0.4
	LAST	0.1	-0.1	0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.8	-1.2	-0.0	-0.3
	.400000	0.5	0.6	0.6	-0.9
	.600000	-0.1	0.3	0.2	0.0
	.800000	0.0	-0.1	-0.7	0.5
	LAST	-1.0	0.7	-0.8	0.8
YES	.200000	-0.8	1.2	0.2	0.4
	.400000	-0.4	-0.7	-0.7	1.0
	.600000	0.4	-0.0	0.1	0.0
	.800000	0.2	0.4	0.6	-0.3
	LAST	0.2	-0.0	0.1	-0.1

PAGE 17 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.8	2.3	-1.4	-1.3
	.400000	1.1	-1.1	0.0	0.6
	.600000	0.3	-0.9	0.5	0.5
	.800000	0.1	0.3	-0.7	-0.7
	LAST	-1.0	0.8	-0.5	0.5
YES	.200000	1.9	-2.4	1.5	1.4
	.400000	-1.2	1.2	0.0	-0.7
	.600000	-0.3	1.0	-0.6	-0.6
	.800000	-0.1	-0.3	0.6	0.6
	LAST	0.1	-0.1	0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-2.0	2.1	-1.6	-1.4
	.400000	1.0	-1.1	0.0	0.6
	.600000	0.4	-0.8	0.6	0.6
	.800000	0.3	0.4	-0.7	-0.7
	LAST	-1.2	0.8	-0.5	0.6
YES	.200000	1.7	-2.7	1.3	1.3
	.400000	-1.3	1.1	0.0	-0.7
	.600000	-0.2	1.0	-0.6	-0.6
	.800000	0.1	-0.2	0.6	0.6
	LAST	0.2	-0.1	0.4	-0.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOCP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	-1.3	-0.1	0.0	0.0	0.0	2.9	-1.1	-0.8	0.0	-1.1	-0.8
	.400000	0.0	-0.7	1.2	0.0	0.0	0.0	-0.9	0.0	0.0	0.0	0.0	0.0
	.600000	-0.8	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.1	0.6	-0.7	0.0	0.1	0.0	0.5	0.0	0.0	0.0	0.0	0.0
	LAST	1.2	-0.3	-0.1	-0.2	-0.4	-0.3	-0.5	-0.1	-0.1	-0.3	-0.2	-0.3
YES	.200000	0.0	1.6	0.1	0.0	0.0	0.0	-3.4	1.3	0.9	0.0	-1.3	0.9
	.400000	0.0	0.8	-1.3	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.9	0.0	-0.6	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.800000	-0.1	-0.5	0.6	0.0	-0.1	0.0	-0.5	0.0	0.0	0.0	0.0	0.0
	LAST	-0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0

OUT	LOCP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	0.4	-1.5	0.0	0.0	-0.8	-1.1	0.0	-0.8	-0.8	-0.8	0.5	0.5
	.400000	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.5	-0.8	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	-0.7	0.0	0.0	0.0	0.0
	LAST	-0.1	2.0	-0.1	-0.4	-0.2	1.8	-0.1	-0.2	-0.4	-0.1	-0.2	-0.1
YES	.200000	-0.5	1.8	0.0	0.0	0.9	1.3	0.0	0.9	0.9	0.9	-0.6	-0.6
	.400000	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	-0.6	0.9	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.6	0.0	0.0	0.0	0.0
	LAST	0.0	-0.3	0.0	0.1	0.0	-0.2	0.0	0.0	0.1	0.0	0.0	0.0

OUT	LOCP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	.200000	0.0	-0.2	-0.6	0.5	0.0	-0.9	0.0	-0.2	0.0	-1.1	0.0	0.5
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	-0.3	-0.3	-0.2	0.0	-0.1	-0.5	-0.3	-0.2	-0.3	-0.1	-0.2	-0.3
YES	.200000	0.0	0.2	0.7	-0.6	0.0	1.0	0.0	0.2	0.0	1.3	0.0	-0.6
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0

OUT	LOCP	BRN 47.0000	48.0000
NO	.200000	0.5	-0.8
	.400000	0.0	0.0
	.600000	-0.2	0.0
	.800000	0.0	0.0
	LAST	-0.1	-0.2
YES	.200000	-0.6	0.9
	.400000	0.0	0.0
	.600000	0.2	0.0
	.800000	0.0	0.0
	LAST	0.0	0.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOOP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.2	0.2	-0.4	0.2	0.1	-0.5	0.0	0.1	0.0	0.0	0.1	0.1
	.400000	0.6	0.1	-0.8	1.0	0.0	-1.1	0.0	-0.7	0.0	0.6	0.6	0.0
	.600000	0.5	0.5	0.0	-1.3	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5
	.800000	0.0	-0.0	-0.8	0.0	0.1	-0.7	0.0	0.0	0.8	0.8	0.0	0.8
	LAST	-0.3	-0.9	0.0	0.8	-0.5	4.2	-0.3	-0.3	-0.1	-0.2	0.0	-0.3
YES	.200000	-1.0	-0.6	1.6	-0.6	-0.5	2.0	0.0	-0.3	0.0	0.0	-0.2	-0.2
	.400000	-0.7	-0.1	0.9	-1.2	0.0	1.2	0.0	0.8	0.0	-0.7	-0.7	0.0
	.600000	-0.6	-0.6	-0.1	1.6	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	-0.6
	.800000	0.0	0.0	0.8	0.0	-0.1	0.6	0.0	0.0	-0.7	-0.7	0.0	-0.7
	LAST	0.0	0.1	-0.0	-0.1	0.1	-0.6	0.0	0.0	0.0	0.0	0.0	0.0

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.3	0.2	-0.3	0.2	0.2	-0.4	0.0	0.2	0.0	0.0	0.2	0.2
	.400000	0.6	0.3	-0.7	1.0	0.0	-1.3	0.0	-0.8	0.0	0.6	0.6	0.0
	.600000	0.6	0.6	0.2	-1.8	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	0.1	-0.7	0.0	0.3	-0.7	0.0	0.0	0.7	0.7	0.0	0.7
	LAST	-0.2	-1.1	0.2	0.7	-0.5	1.3	-0.1	-0.1	-0.0	-0.1	0.0	-0.2
YES	.200000	-1.2	-0.6	1.2	-0.6	-0.4	1.1	0.0	-0.2	0.0	0.0	-0.1	-0.1
	.400000	-0.7	0.0	0.9	-1.5	0.0	1.0	0.0	0.7	0.0	-0.7	-0.7	0.0
	.600000	-0.6	-0.6	0.1	1.3	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	-0.6
	.800000	0.0	0.1	0.8	0.0	0.1	0.6	0.0	0.0	-0.8	-0.8	0.0	-0.8
	LAST	0.1	0.2	0.0	-0.1	0.1	-0.4	0.1	0.1	0.2	0.2	0.0	0.1

PAGE 32 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOGP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.2	-0.2	0.3	0.8
	.400000	-0.3	0.7	0.0	-0.7
	.600000	0.2	0.2	0.0	-0.8
	.800000	-0.9	0.5	1.4	-0.7
	LAST	-0.8	-0.9	-0.4	2.4
YES	.200000	0.2	0.2	-0.[illegible]	-0.9
	.400000	0.3	-0.8	0.0	0.8
	.600000	-0.2	-0.2	0.0	0.9
	.800000	0.8	-0.5	-1.3	0.6
	LAST	0.1	-0.1	0.1	-0.3

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.2	-0.1	0.4	0.8
	.400000	-0.2	0.8	0.0	-0.8
	.600000	0.3	0.3	0.0	-0.8
	.800000	-0.8	0.6	1.2	-0.7
	LAST	-0.8	0.8	-0.3	1.2
YES	.200000	0.2	0.3	-0.2	-0.8
	.400000	0.4	-0.7	0.0	0.7
	.600000	-0.1	-0.0	0.0	0.8
	.800000	0.8	-0.3	-1.8	0.6
	LAST	0.1	-0.1	0.1	-0.2

PAGE 14 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.2	-1.0	0.0	0.0
	.400000	-0.2	0.3	0.3	0.1
	.600000	0.3	-0.2	0.8	-0.5
	.800000	-0.6	0.0	0.8	0.1
	LAST	-1.3	1.0	-1.2	1.2
YES	.200000	-1.0	5.6	0.0	0.0
	.400000	0.3	-0.5	-0.5	-0.1
	.600000	-0.4	0.2	-0.9	0.6
	.800000	0.3	-0.0	-0.5	-0.1
	LAST	0.3	-0.2	0.3	-0.2

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.2	-1.2	0.0	0.0
	.400000	-0.1	0.5	0.5	0.2
	.600000	0.4	-0.0	0.8	-0.4
	.800000	-0.5	0.1	0.8	0.2
	LAST	-1.7	0.9	-1.6	1.1
YES	.200000	-1.2	1.4	0.0	0.0
	.400000	0.4	-0.5	-0.5	0.1
	.600000	-0.2	0.3	-1.1	0.6
	.800000	0.4	0.1	-0.4	0.0
	LAST	0.3	-0.2	0.3	-0.2

PAGE 19 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.0	0.0	0.0	0.0
	.400000	0.1	0.1	0.0	-0.4
	.600000	0.2	-0.4	-0.1	0.8
	.800000	1.0	-0.3	-0.7	-0.5
	LAST	-0.9	0.5	0.3	0.3
YES	.200000	0.0	0.0	0.0	0.0
	.400000	-0.2	-0.1	0.0	0.6
	.600000	-0.3	0.4	0.2	-0.9
	.800000	-0.6	0.2	0.4	0.3
	LAST	0.2	-0.1	-0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.0	0.0	0.0	0.0
	.400000	0.2	0.1	0.0	-0.2
	.600000	0.3	-0.3	0.1	0.8
	.800000	1.0	-0.3	-0.8	-0.4
	LAST	-0.8	0.6	0.4	0.4
YES	.200000	0.0	0.2	0.0	0.0
	.400000	-0.0	0.0	0.0	0.6
	.600000	-0.1	0.5	0.3	-1.1
	.800000	-0.5	0.2	-0.5	-0.4
	LAST	0.2	-0.1	-0.0	-0.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	0.0	-0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.2	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.5	0.0	0.0	0.0	-0.9	0.0	0.0	0.0	0.0	0.0
	.800000	0.1	0.9	-0.5	-0.5	-0.7	-0.5	0.3	0.0	0.0	-0.5	0.0	0.0
	LAST	0.2	-0.5	0.0	-0.2	1.1	-0.4	-0.4	-0.2	-0.2	-0.5	-0.3	-0.5
YES	.200000	0.0	0.0	1.5	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	-0.4	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	-0.5	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0
	.800000	-0.1	-0.5	0.3	0.3	-0.4	0.3	-0.2	0.0	0.0	0.3	0.0	0.0
	LAST	-0.1	0.1	0.0	0.0	-0.2	0.1	0.1	0.0	0.0	0.1	0.1	0.1

OUT	LRPP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	-0.5	-0.5	0.0	0.0	-0.1	-0.5	1.4	0.0	0.0	-0.5	0.6
	LAST	-0.2	0.8	0.0	-0.7	-0.4	0.6	0.0	-0.4	-0.7	-0.2	2.4	0.0
YES	.200000	-0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.3	0.3	0.0	0.0	0.0	0.3	-0.9	0.0	0.0	0.3	-0.4
	LAST	0.0	-0.2	0.0	0.1	0.1	-0.1	0.0	0.1	0.1	0.0	-0.5	0.0

OUT	LRPP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.6	0.0	-0.5	-0.5	0.0	0.6	0.0	0.0	0.0	-0.5
	LAST	-0.4	-0.5	-0.3	0.0	0.0	0.6	-0.5	-0.3	-0.5	-0.2	-0.3	1.4
YES	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	-0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	-0.4	0.0	0.3	0.3	0.0	-0.4	0.0	0.0	0.0	0.3
	LAST	0.1	0.1	0.1	0.0	0.0	-0.1	0.1	0.1	0.1	0.0	0.1	-0.3

OUT	LRPP	BRN 47.0000	48.0000
NO	.200000	0.0	0.0
	.400000	0.0	0.0
	.600000	0.2	0.0
	.800000	-0.5	0.0
	LAST	0.0	-0.3
YES	.200000	0.0	0.0
	.400000	0.0	0.0
	.600000	-0.3	0.0
	.800000	0.3	0.0
	LAST	0.0	0.1

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	AMTIN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.1	0.0	-0.2	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.6	-0.1	-0.5	0.6	0.0	-1.2	0.0	0.3	0.0	0.0	0.3	0.3
	.600000	0.6	0.4	-1.1	0.2	0.0	-0.8	0.0	0.0	0.0	0.6	0.0	0.6
	.800000	-0.5	-0.1	-0.1	-1.0	1.9	-0.7	0.0	-0.5	1.4	0.0	0.0	0.6
	LAST	1.4	-1.4	-0.8	1.1	-0.9	5.1	-0.4	-0.4	-0.2	2.4	0.0	-0.5
YES	.200000	-0.6	-0.2	1.4	-0.3	-0.3	-0.2	0.0	-0.2	0.0	0.0	-0.2	0.0
	.400000	-0.9	0.2	0.8	-0.9	0.0	1.9	0.0	-0.5	0.0	0.0	-0.5	-0.5
	.600000	-0.7	-0.4	1.3	-0.3	0.0	0.9	0.0	0.0	0.0	-0.7	0.0	-0.7
	.800000	0.3	0.1	0.1	0.6	-1.1	0.4	0.0	0.3	-0.9	0.0	0.0	-0.4
	LAST	-0.3	0.3	0.2	-0.2	0.2	-1.1	0.1	0.1	0.0	-0.5	0.0	0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	AMTIN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.2	0.2	-0.2	0.2	0.2	0.2	0.0	0.2	0.0	0.0	0.2	0.0
	.400000	0.6	-0.0	-0.4	0.6	0.0	-1.6	0.0	0.5	0.0	0.0	0.5	0.5
	.600000	0.6	0.5	-1.1	0.3	0.0	-0.8	0.0	0.0	0.0	0.6	0.0	0.6
	.800000	-0.4	0.0	-0.0	-1.0	1.4	-0.8	0.0	-0.4	1.0	0.0	0.0	0.6
	LAST	1.0	-2.0	-0.7	1.0	-1.0	1.9	-0.3	-0.3	-0.1	1.2	0.0	-0.4
YES	.200000	-0.6	-0.1	1.0	-0.2	-0.2	-0.1	0.0	-0.1	0.0	0.0	-0.1	[illegible]
	.400000	-1.1	-0.3	0.8	-1.1	0.0	1.3	0.0	-0.5	0.0	0.0	-0.5	[illegible]
	.600000	-0.6	-0.3	1.1	-0.1	0.0	0.8	0.0	0.0	0.0	-0.6	0.0	-0.6
	.800000	0.4	0.1	0.1	0.6	-1.2	0.5	0.0	0.4	-1.0	0.0	0.0	-0.2
	LAST	-0.2	0.3	0.2	-0.2	0.2	-1.1	0.2	0.2	0.2	-0.4	0.0	0.2

PAGE 35 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.0	0.1	0.0	0.0
	.400000	0.2	-0.4	0.0	0.3
	.600000	-0.2	0.5	0.0	0.0
	.800000	0.2	-0.1	0.1	-0.7
	LAST	-1.5	2.4	1.1	1.0
YES	.200000	0.2	-0.3	-0.2	-0.2
	.400000	-0.4	0.7	0.0	-0.5
	.600000	0.3	-0.5	0.0	0.0
	.800000	-0.1	0.1	-0.1	0.4
	LAST	0.3	-0.5	-0.2	-0.2

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	AMT 250.000	500.000	750.000	LAST
NO	.200000	0.0	0.2	0.2	0.2
	.400000	0.3	-0.3	0.0	0.5
	.600000	-0.2	0.5	0.0	0.0
	.800000	0.3	0.0	0.2	-0.8
	LAST	-1.6	1.6	0.9	0.8
YES	.200000	0.4	-0.2	-0.1	-0.1
	.400000	-0.3	0.7	0.0	-0.5
	.600000	0.3	-0.4	0.0	0.0
	.800000	-0.1	0.1	0.0	0.5
	LAST	0.3	-0.5	-0.1	-0.1

The NEW variable should also be examined. Here, there are three cells with standardized deviates greater than two (the largest is 2.6). The number of observed cases in these cells totalled 17. Once again, a total of only 17 cases (a small percentage: 17/339 = 5.01 percent) accounts for the large Pearson chi-square value.

The two-factor interaction term for the BRN variable was not significant, which implies a good fit of the data is accommodated by the model.

The TY*NEW variable had 2 cells with standardized deviates greater than 2.0 (with values of 2.6 and 2.2). The number of observations in these cells totalled 6. Again, only small cells are involved.

The LRFE Model

For the fixed effects logit model, the NEW variable had a significant two-factor interaction. An examination of the standardized deviates shows two cells with values greater than 2 (values of 2.3 and -2.4) in category two. The number of observations in these cells totalled 62. Category two relates to new firms with no previous bank history. The expected values underestimated the value for the no group and similarly overestimated the yes group for category two of the NEW variable. As a result, firms in this category were, for the most part, denied loans even though these firms may have scored high on other variables. Hence, yes responses were expected but no's were observed. In all likelihood, this is due to the unwillingness of the bank to take risks with "new" customers.

The AMT/TNW (or AMTTN) also had significant two- factor interaction. This variable had one cell with a large standardized deviate of 4.2. However, the number of observed cases in that cell was only one (with an expected value of 0.1).

The LRME Model

The TY variable for the logit model with mixed effects had only one cell with a large standardized deviate (5.6). This originates from a cell with an expected value of near zero and an observed frequency of one.

Finally, the AMTTN had a single large standardized deviate of 5.1. This originated from a cell with 2 cases and an expected value of 0.1.

Conclusions

The objective of this residual analysis was to determine the aptness of the models presented in Section 5.1. Detailed plots and histograms failed to show any underlying pattern in the residuals with respect to either the "yes" or "no" groups. The chi-square tests prove to be significant but due to the small number of observations in many of the branches, the tests for the significance of the two-factor interaction were inconclusive. In this exercise, we did not determine any departures from the fit of the logistics models.

Examination of the plots did reveal that some problems with the ratio

AMT/TNW might explain the counter-intuitive result of the coefficient (positive instead of negative). As stated above, the presence of an outlier for AMT/TNW was detected in this residual analysis. Model fitting was carried out both with and without this one observation, and the results were fairly stable. Consequently, this observation was left in the data for the remainder of the model analysis.

In conclusion, all this implies a similarity of distributions of independent variables, given a level of p, for both the y=1 and y=0 groups. According to Rubin (1984), the models then provide a good fit to the data. In short, we can state that there is no evidence of a lack of appropriateness in the models presented here.

Comparison of the three models can also be undertaken. From the histograms and the log-linear tables, it can be seen that LRME provides the best fit to the data with LDA and LRFE not performing as well. The LRME, for instance, has only two of the three-factor interactions significant. Further, the two-factor interactions contain only two which are significant. Both the LDA and LRFE models contain more significant interactions for both the two- and three-factor tests. Although, intuitively one would expect the fixed effects model to fit better, the residual analysis indicates the random effects model does well and this may be due to the fact that that are so many cells with few observations.

Consequently, we now proceed with tests of effectiveness on the models.

5.3 Classification Results

In the preceding two sections of this chapter, we discussed the goodness of fit of the "best" models for the three different statistical techniques. Stepwise procedures and residual analysis were discussed; however, the primary goal of this research is to establish the most effective method for classification. The ability of the models to classify the cases in the data (the same cases that were used in the model building) is one measure of their effectiveness. To this end, the three methods (LDA, LRFE and LRME) were applied to the original data set and then they were tested by classifying each of the cases in this set. Table A-7 contains classification results for the original data set. A comparison below focusses on an overall classification (yes and no groups combined) as well as individual classifications (yes and no group separately).

In this study, the no's are more difficult to classify than the yes observations. This is due to the fact that there are less no's in the sample. Hence, groups with smaller probabilities are more difficult to classify. The performance of the different statistical methods in classifying the no's is charted in Table A-7.

Overall, the LRFE model performs better than the other two methods; the percentage of cases classified correctly is 91.87 percent. It is expected, when classifying the original data set (using the same data that was used in the model building), that the LRFE would perform better than the LRME because the LRFE model contains more parameters. This enables the LRFE to classify the original data set more effectively. In fact, we could increase the number of

Table A-7
Results for Best Models

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	75.86	66	21	87
	Yes	91.27	22	230	252
	Total	87.32	88	251	339
	% correct		75.00	91.63	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	89.87	71	8	79
	Yes	92.65	15	189	204
	Total	91.87	86	197	283
	% correct		82.56	95.94	
C.	Logit Regression (Mixed Effects) (LRME)				
	No	72.15	57	22	79
	Yes	94.12	12	192	204
	Total	87.99	69	214	283
	% correct		82.61	89.72	

Table A-7.1
Best Model LDA on sample of 283 cases

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	70.89	56	23	79
	Yes	94.61	11	193	204
	Total	87.99	67	216	283
	% correct		83.58	89.35	
B.	LDA* (adjusted for cases and effects)				
	No	78.48	62	17	79
	Yes	91.18	18	186	204
	Total	87.63	80	203	283
	% correct		77.50	91.63	

Table A-7.2
Comparison of Naive Models

	Observed	% correct	Classified No	Classified Yes	Total
Naive Model					
	No	0.00	0	87	87
	Yes	100.00	0	252	252
	Total	74.34	0	339	339
	% correct			74.34	
Challenging Naive Model					
	No	43.67	38	49	87
	Yes	89.29	27	225	252
	Total	77.58	65	274	339
	% correct		58.46	82.12	
Special Branch Naive Model					
	No	77.01	67	20	87
	Yes	84.52	39	213	252
	Total	82.60	106	233	339
	% correct		63.21	91.42	
Naive Model Based on Branches					
	No	78.16	68	19	87
	Yes	84.52	39	213	252
	Total	82.89	107	232	339
	% correct		63.55	91.81	

parameters to equal the number of observations in the data, and as a result, would obtain 100 percent accurate classification. The problem with this methodology, of course, is that such a model would overfit the data, and cause inaccurate results when predicting cases that were not used in the model building process. The LRME model does marginally better than the LDA with respect to overall classification (87.99 percent compared to 87.32 percent).

Upon examination of the "no" group, we see that each method performs less effectively than it did for the "yes" group. Once again, the LRFE is the best method. The LDA performs slightly better than the LRME in percentage correct for the "no" group, and for the "yes" group, the LRME performs slightly better than the other two models.

It should be noted at this point, that the number of cases used in the evaluation of the logit model is less than the 339 cases of the original data set used in the evaluation of the LDA. This disparity resulted because the TNW variable which is included in the LRFE and LRME models, was missing in 56 of the cases. Thus, the LRFE and LRME models were fit on the remaining 283 cases. In this section, the models are evaluated on as many cases as possible (339 for LDA, and 283 for LRFE and LRME). In Section 5.7 of this chapter, we investigate other techniques for handling the problem of missing data.

In order to compare the three methods on equal ground, the LDA model (not LDA*) was run on the 283 cases that were used for the logit models. These classification results are given in Table A-7.1. As can be seen, the overall classification is almost exactly the same on this subset of the data as it was on the full sample (87.99 percent compared to 87.32 percent correct). The yes and no groups differ slightly in percentage classified correctly. Since the smaller data set supplies little difference, the remaining analysis for LDA is performed on the full 339 cases. As well, the adjusted LDA method is applied throughout, LDA*, and it performs, overall, comparably to the original LDA method at this step.

5.4 Naive Models

As a standard for comparison, we consider a "naive" model where all the cases are classified into the modal group. If this ~~type of model performs better~~ than the more sophisticated techniques presented in this research, then the need for these sophisticated methods should be questioned. The results of the naive model can be seen in Table A-7.2. Since the majority of cases are yes, all the cases will be classified as yes, leading to an overall effectiveness of 74.34 percent. Inspection of the yes and no groups, respectively, shows a 100 percent and 0 percent correct classification.

As a more serious comparison, consider a "challenging naive" model where cases are classified into the modal group with cells derived from a couple of selected variables. As described in Section 5.1, three prominent variables across all 3 techniques were amount of the loan (AMT), type of the loan (TY), and experience of firm (NEW). The amount of the loan was made discrete and was calculated from a variable with 4 possible values:

$ 0 - 250,000 for group 1

250,001 - 500,000 for group 2

500,001 - 750,000 for group 3

750,001 - 1,000,000 for group 4

Once again, TY and NEW both have 4 categories. This then produced a table with 64 categories (4x4x4). The results of this "challenging naive" classification scheme are given in Table A-7.2. The overall classification success rate climbs approximately 3 percent to a total of 77.58 percent. This increase in effectiveness was due to the no's outnumbering the yes's in two of the cells. We see that this classification methodology performs better for the hard to predict no group than did the simple naive model for the same group (43.67 percent correct as opposed to 0 percent). Although the "challenging naive" model performs better than the "naive" model, neither of these techniques approach the effectiveness of the more sophisticated models.

We can consider another naive model based on the branches, the Special Branch Naive Model classification rule (SBNM), where any case observed from branch 4, 11, or 19 is classified as a no and all the other cases are classified as a yes. This type of a naive model classifies 82.60 percent correct.

As a last case, let us consider a naive model where classification is based on the modal group within each branch. These results are also given in Table A-7.2 under "Naive Model based on Branches" (NMBB).

	Overall	No Group	Yes Group
LDA -	87.32	75.86	91.27
LRFE -	91.87	89.87	92.65
LRME -	87.99	72.15	94.12
LDA* -	87.63	78.48	91.18
Naive -	74.34	0.00	100.00
Challenging			
naive -	77.58	43.67	89.29
SBNM -	82.60	77.01	84.52
NMBB -	82.89	78.16	84.52

With respect to overall classification, the best model is the LRFE model. In fact this model is approximately 4 percentage points better than any of its competitors. The other "sophisticated" models, LDA (LDA*) and LRME, are grouped together. The naive models are at least five percentage points in effectiveness below the "sophisticated" models.

Do our more sophisticated models perform that much better than this naive model? This question leads directly to an investigation of the more advanced modelling methods described in Sections 5.8 and 5.9.

5.5 Results for the Holdout Samples

Classification effectiveness is an important measure for evaluation of the LDA, LRFE and LRME models. However, classification solely involves the grouping of the same data that were used in fitting the models. An alternate measure of effectiveness is the model's ability to predict new cases. This involves randomly splitting the original data for the purpose of testing predictive ability on holdout samples (see Section 4.4). First, the data were split in half and each model (LDA, LRFE and LRME) was fit on the first half; the fitted model was then tested on the second half. The results on the first half (the training sample) yield the classification results, and the results on the second half (the holdout sample) represent the percentage correct in predicting new cases. This procedure was performed five times.

Second, the original data was split in an 80-20 manner where 80 percent of the data was used to fit the model and 20 percent was used to measure the predictive ability. This procedure was also performed five times. Therefore, in total, ten different holdout samples were tested. The same splits were used for each of the three statistical methods, and the results are presented in Appendix A, Tables A-8. This allows for a direct comparison of the three models. The LDA* model was also evaluated according to the same procedure, but the samples were different from those used for the other models. This difference is indicated by the separation of LDA* in the tables in the appendix. In Table A-8, the results for the training samples are given and are numbered from one to ten. In Table A-8.1, the results for the holdout samples are presented and the same numbering sequence is followed; this sequence illustrates the ten replications of holdout analysis.

Overall results are listed below in Table 5.5.1. Here, the mean and standard deviation of the percentage of correctly classified cases are given for both splits of the data and for the overall data. In the training sample, the LRFE performs better than the other two methods for both types of splits (50-50 and 80-20). Consequently, if concentrate on the classification of the training set, the LRFE (93.539 percent correctly classified) would seem to be the best overall model. (This is the same result we saw in Section 5.3.)

However, a better evaluation technique to measure the models' effectiveness would be to examine prediction results. In the second half of Table 5.5.1, the LRME out-performs the other two methods in the three categories - the two splits and overall. On average, 84.298 percent of the cases in the holdout groups are classified correctly by this method. This number exceeds the percentage correct for LDA and greatly exceeds that attained by LRFE and LDA*. In fact, a paired difference test comparing the percentage correct for the holdout samples for the LRME and the LDA shows that the LRME performs statistically significantly better ($p < .01$).

Detailed analysis was also performed for the yes groups and the no groups separately. These results are presented in Table 5.5.2. For the yes group, the LRFE model performs better for classification in the training sample, and the LRME model has the highest percentage correct in predicting new cases in the holdout sample. These results are consistent with those produced in the overall analysis.

TABLE 5.5.1

Holdout Analysis

	LDA	LRFE	LRME	LDA*
50 - 50 split:				
Classification of Training sample				
Mean	86.972	94.724	87.596	93.034
Standard Deviation	1.603	2.217	2.533	1.411
Prediction of Holdout sample				
Mean	81.032	77.236	84.296	76.504
Standard Deviation	1.682	4.015	1.428	1.037
80-20 split:				
Classification of Training sample				
Mean	86.452	92.354	89.334	89.184
Standard Deviation	1.284	1.212	1.301	1.281
Prediction of Holdout sample				
Mean	80.720	78.060	84.300	76.306
Standard Deviation	2.633	4.271	3.963	4.526
Overall Results:				
Classification of Training sample				
Mean	86.712	93.539	88.465	91.109
Standard Deviation	1.396	2.097	2.108	2.394
Prediction of Holdout sample				
Mean	80.876	77.648	84.298	76.405
Standard Deviation	2.089	3.973	2.798	3.097

TABLE 5.5.2

Holdout Analysis - YES Group

	LDA	LRFE	LRME	LDA*
50 - 50 split:				
Classification of Training sample				
Mean	89.166	96.178	93.268	95.030
Standard Deviation	2.482	1.438	4.045	1.803
Prediction of Holdout sample				
Mean	88.502	84.038	91.536	82.254
Standard Deviation	5.652	7.197	4.669	4.929
80-20 split:				
Classification of Training sample				
Mean	88.922	93.398	91.788	91.462
Standard Deviation	3.171	2.199	2.445	1.696
Prediction of Holdout sample				
Mean	85.558	81.726	94.048	83.950
Standard Deviation	5.771	8.566	3.019	3.583
Overall Results:				
Classification of Training sample				
Mean	89.044	94.788	92.528	93.246
Standard Deviation	2.687	2.284	3.246	2.502
Prediction of Holdout sample				
Mean	87.003	82.882	92.792	83.102
Standard Deviation	5.644	7.558	3.936	4.159

TABLE 5.5.2

Holdout Analysis – NO Group

	LDA	LRFE	LRME	LDA*
50 – 50 split:				
Classification of Training sample				
Mean	80.758	90.614	72.484	88.062
Standard Deviation	3.125	7.022	8.757	6.095
Prediction of Holdout sample				
Mean	59.042	59.404	65.176	62.162
Standard Deviation	14.114	6.881	13.606	10.783
80-20 split:				
Classification of Training sample				
Mean	79.746	89.714	83.242	83.280
Standard Deviation	5.424	5.368	2.862	4.135
Prediction of Holdout sample				
Mean	61.796	66.298	55.274	55.962
Standard Deviation	20.948	15.534	12.632	14.194
Overall Results:				
Classification of Training sample				
Mean	80.252	90.164	77.863	85.671
Standard Deviation	4.207	5.911	8.359	5.519
Prediction of Holdout sample				
Mean	60.419	62.851	60.225	59.061
Standard Deviation	16.902	11.895	13.433	12.324

For the no group, the percentage correct values for the holdout groups are very close. In order to see if the difference was significant across methods, analysis of variance based on a randomized block design was performed on the holdout sample results for all methods except LDA*. The results from this test indicate that there are no significant differences in the mean values for any of the three methods ($p > 0.200$). Thus, even though LRFE performs slightly better in the no group, this difference is not significant. The LDA* model was not included in the significance test due to the different samples involved, but it can be observed that this model has a lower percentage correct than the other methods.

The results of the holdout analysis concur with the classification results on the original data. In testing the same data that was used in fitting the model, the LRFE performs the best. The LDA* model also performs well; but it should be noted that the LRFE, LDA, and LDA* all contain more parameters than the LRME model. Consequently, these models have a tendency to overfit the data for prediction. When a new data set (a holdout sample) is used, then a random effects model is better (the LRME model is the best). Efron and Morris (1972, 1973 and 1975), illustrate that the empirical Bayes or shrinkage estimates perform better than fixed effects methods.

5.6 Results for the Bootstrap Methods

A second methodology for measuring prediction ability is the bootstrap method described in Chapter Four. Again, emphasis is placed on the classification of new cases rather than on the cases used in the fitting of the model. In order to obtain a reliable estimate of the true prediction error rate, the bootstrap methodology was repeated 20 times for each of the three models. More specifically, the first "resample" was created and was fitted to the "best" model for LDA, LRFE and LRME. These fitted models were then used to classify the data in the original data set. Through procedures described in detail in Chapter Four, an estimate of the expected excess error (EEE) rate was determined. This procedure was repeated 20 times in order to produce both a more reliable estimate and a good estimate of the variability. The results for the 20 bootstrap samples are given in the appendix. Table A-9 lists the results of the classification of the data that were used in the model fitting. These tables are numbered from one to twenty; the numbers represent the different replications. Table A-9.1 gives the results of the ability of the fitted models (for each of the twenty replications) to predict the data from the original data set. Once again, the numbering system is the same for both the classification and the prediction results; this allows for a direct comparison of the information. (Again, the LDA* results are separated in the tables.)

Table 5.6.1, below, summarizes the overall results for EEE. As can be seen, the LDA method performs slightly better than the LRME; and the LDA, the LDA* and the LRME do substantially better than the LRFE in EEE.

The results from the samples show a statistically significant difference ($p<.005$) for a "paired difference test" involving LDA and LRME for EEE. The true error rates can be calculated by adding the EEE value (4.1885 for LDA) to the error rate of the model when classifying the original data set (12.68 for

TABLE 5.6.1

Bootstrap Analysis

Replication Number	EEE			
	LDA	LRFE	LRME	LDA*
1	7.37	8.79	7.42	6.71
2	2.07	14.75	7.93	1.41
3	2.65	9.33	4.44	4.24
4	3.25	6.85	4.44	8.48
5	5.02	13.76	4.92	5.30
6	1.47	10.32	5.75	5.65
7	7.08	14.87	5.43	3.89
8	4.42	7.35	4.85	5.66
9	5.31	10.81	6.88	2.12
10	5.31	18.74	6.03	4.59
11	5.31	13.07	6.01	7.07
12	6.49	12.27	8.18	8.48
13	7.08	11.23	8.28	5.65
14	3.54	6.47	5.12	11.66
15	6.49	9.25	6.09	9.18
16	3.24	5.88	6.86	9.18
17	2.07	14.49	6.01	1.06
18	2.06	15.37	6.60	4.94
19	2.36	7.14	3.63	2.48
20	1.18	9.67	5.46	5.66
Mean	4.1885	11.0205	6.0165	5.6705
Standard Deviation	2.0496	3.5464	1.2977	2.7965

LDA; calculated by taking the percentage correct - 87.32 from Table A-7 - and subtracting it from 100; this gives the final result of 16.8685). The resulting estimates of the "true error rates" are as follows:

LDA = 16.8685

LRFE = 19.1505

LRME = 18.0265

LDA* = 18.0405 .

Similarly, the percentage correct based on this error rates can be easily calculated, and they are given below for completeness.

LDA = 83.1315

LRFE = 80.8495

LRME = 81.9735

LDA* = 81.9595 .

These results once again support the hypothesis that the LRFE does poorly in its prediction relative to its ability to classify the same data used in fitting the model. As well, an initial interpretation would lead to a conclusion that LDA out-performs LRME by a significant margin. However, LDA has been calculated on a larger sample and a more accurate comparison would be based on LDA* versus LRME. At this point, it would appear that LRME performs marginally better. (Remember that we are interested in true error rates, not just EEE.)

However, to further compare LDA with LRME, a "paired difference" test was performed on the percentage predicted correct for each of the 20 replications of the bootstrap. The percentage predicted correct, is determined by using the model developed on the resample to predict the cases in the original data set. This procedure provides another measure of the effectiveness of these two methods. Here, as it can be seen, the LRME was significantly better than the LDA ($p < 0.0001$) (see Table 5.6.2).

This result clearly illustrates that the LRME performs better than the LDA in prediction and is no worse (and probably better) in true error rate over LDA*. The fact that only 20 replications were executed may be one of the reasons why the results of the bootstrap were so close across methods. One result which is clear is that the LRME methods was not out-performed by any method.

5.7 Missing Data

In many statistical applications, the data are not complete. One way of addressing this problem is to abandon the cases in which some of the data are missing and use the remaining part of the data set for the analysis. In the analysis presented thus far; only information on total net worth is missing. As this variable was not included in the LDA model, the discriminant analysis included all 339 cases during the model building process. For the logit regression, for both LRFE and LRME, TNW was an integral part of the analysis; therefore, in the 56 cases where information is missing, the cases were removed

TABLE 5.6.2

Percentage Predicted Correct in the Bootstrap Replications

	LDA	LRME	Difference
1.	90.56	92.23	+1.67
2.	87.91	89.56	+1.65
3.	87.02	90.66	+3.58
4.	89.68	88.19	-1.49
5.	90.27	92.55	+2.28
6.	87.61	91.26	+3.65
7.	90.27	90.24	-0.03
8.	88.20	91.07	+2.87
9.	88.79	89.57	+0.78
10.	90.56	92.25	+1.69
11.	90.27	92.23	+1.96
12.	92.33	94.05	+1.72
13.	90.56	92.73	+2.17
14.	87.02	90.28	+3.26
15.	91.45	91.96	+0.51
16.	87.61	88.13	+0.52
17.	87.91	91.17	+3.26
18.	87.61	92.47	+4.86
19.	88.20	91.26	+3.06
20.	88.20	90.97	+2.77

and the model building subsequently incorporated the remaining 283 cases.

An alternative method for handling the missing data would be to estimate the missing values for TNW. This can be done using the EM algorithm described earlier (Dempster, Laird and Rubin, 1977). The first step of the estimation of the missing data is to set an initial estimate for the missing values of TNW. In this case, we used the overall mean for the TNW variable. Then a linear regression model is formed where the independent variables are the variables in the logit model where the data are complete; TNW is the dependent variable. The regression coefficients are determined and are used along with the independent variables to predict new values for the missing cases. At this point, a second regression model is developed using these new predicted values as the estimates for the missing values. This procedure continues cycling back and forth until the estimates converge.

Consider the vector

$$\bar{y} = \begin{bmatrix} y \\ \hat{y} \end{bmatrix}. \qquad (5.7.1)$$

where $\hat{y}$ represents the values for TNW where the observations are missing, and y represents the observed values. In the E-step, estimates for the missing values are employed and the vector of the regression coefficients, b, is estimated

$$b = (X'X)^{-1} X' \bar{y} \tag{5.7.2}$$

where X is the matrix of observed values for all the other variables in the model, and $\bar{y}$ is the vector of all TNW values - both the observed and the estimated. Once **b** has been calculated (M-step), the vector $\bar{y}$ is updated by calculating the updated estimates for $\hat{y}$ (call it $\hat{y}_1$)

$$\hat{y}_1 = X\,b \; . \tag{5.7.3}$$

We then proceed back to the E step with the new values for $\hat{y}$, and the process continues until the values for $\hat{y}$ converge.

The results from this method are very encouraging. The values of TNW converged rather quickly (after ten iterations), and the estimates were sensible with no outliers present.

Once the "complete" data set was formed, the logit regression (both the fixed effects and the mixed effects) models were fit on the data. The EM algorithm was used for the LRME in the same manner as discussed in Sections 4.2 and 5.3.

In Table A-10, the classification results on the new data set, where the missing values were estimated, are provided. Again, LRFE performs better than LRME for both the overall group and for the individual yes and no groups. The branch variability values for the LRME model for σ_1^2 and σ_2^2 were

Table A-10

Missing Data: Models Fit on Complete Data Set

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	87.36	76	11	87
Yes	91.74	20	222	242
Total	90.58	96	233	329
% correct		79.17	95.28	
Logit Regression (Mixed Effects) (LRME)				
No	80.46	70	17	87
Yes	90.08	24	218	242
Total	87.54	94	235	329
% correct		74.47	92.77	

Table A-10.1
Missing Data: Original model used to predict new data set of predicted missing values

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	79.31	69	18	87
Yes	89.68	26	226	252
Total	87.02	95	244	339
% correct		72.63	92.62	
Logit Regression (Mixed Effects) (LRME)				
No	72.41	63	24	87
Yes	92.46	19	233	252
Total	87.32	82	257	339
% correct		76.83	90.66	

determined to be 1.69 and .0000033 as compared to 1.89 and .0000042 for the original data set.

Also, in Table A-10.1, the original fitted models (Section 5.1) were used to predict the "new" complete data set. Once again, in predicting new cases, the random effects model out-performs the fixed effects model. Overall, the LRME classified 87.32 percent correct compared to 87.02 percent for the LRFE.

The listing of the regression coefficients and parameter estimates for the LRFE and LRME models (for the 339 cases) are given in Table A-4.

It should be noted that neither residual analysis nor cross validation techniques were carried out during the "missing data" analysis. These analyses were only performed on the original "pooled" approach presented in Section 5.1.

5.8 Separation Models

In both the logit and the discriminant analysis models, the branch variable is significant. However, this significance is primarily due to the oversampling of cases in four specific branches. In more detail, the cases observed from the branches (the 61 declines; see Section 3.2) were obtained from only 4 branches and were selected by the division management of the bank. This deviation in the sampling scheme affected the variable selection process. The variance of the branch effect in the original model (Section 5.1.3) was calculated to be $\sigma^2 = 1.89$.

In order to investigate this branch effect, we consider an additional term in the logit model - whether or not the branch was selected in the second "branch sample". This addition provides us with a "separation" of the branch effects. The 4 branches which were oversampled are examined in one group, and the remaining 44 branches are examined in a second group. We can now measure the effect of the branches of the oversampled group (the 4 branches) as well as the variability of this effect and, next, we can measure the effect of the second group of branches (the remaining 44) and its variability. This allows us to isolate any branch effect which is solely a product of the sampling design.

First, a fixed effects logit model was fit (using BMDPLR). It included:

$$\text{logit}\,(p_{ijklm}) = b_0 + u_{(i)} + BRN_{j(i)} + TY_k + NEW_l + b_1\,AMT + b_2\,AMT/TNW \quad (5.8.1)$$

where $j = 1, \ldots, n_i$;

$n_1 = 4$; $n_2 = 44$; and

$i = 1$ - oversampled group

$i = 2$ - remaining 44 branches.

Thus $BRN_{j(1)}$ represents the branch effect for the oversampled branches ($j = 1,\ldots,4$), and $BRN_{j(2)}$ represents the effect for the remaining branches ($j = 1,\ldots,44$), and $u_{(i)}$ represents the effect for oversampling.

The fixed effects logit regression was applied to the full model described above. The $BRN_{j(2)}$ terms were not found to be significant and were dropped from the model. As a result, the best model (excluding the effect for the 44 branches) was used to classify the cases in the original data set. The classification results are given in Table A-11.

(The logit model with fixed effects was fit where both terms $BRN_{j(2)}$ and $u_{(i)}$ were forced into the model. The classification results for this model are also given in Table A-11 under the heading LRFE*.)

The interaction term between the branch effect and TNW was removed from the model because the model could not be fit with the original number of parameters using the EM algorithm (for the random effects model). Thus this term was excluded for both the LRFE and the LRME models.

The classification results show that this "separation" model performs less effectively than the initial LRFE model presented in Section 5.1.2. Overall, the classification is 86.57 percent correct, whereas the best model classified 91.87 percent correct. This may be due in part to the removal of the branch effect for the 44 branches as well as the interaction term mentioned above.

The coefficients for the significant variables are given in Table A-12. Also listed in this table are the coefficients for the LRFE model where both the $BRN_{j(2)}$ and $u_{(i)}$ terms were forced into the model. This is represented by LRFE* in the table.

Table A-11

"Separation" Models - Classification

Observed	% correct	Classified No	Classified Yes	Classified Total
Logit Regression (Fixed Effects) (LRFE)				
No	64.56	51	28	79
Yes	95.10	10	194	204
Total	86.57	61	222	283
% correct		83.61	87.39	
Logit Regression (Mixed Effects) (LRME)				
No	65.82	52	27	79
Yes	94.61	11	193	204
Total	86.57	63	220	283
% correct		82.54	87.73	
LRFE* (Model with variables forced in)				
No	68.35	54	25	79
Yes	94.61	11	193	204
Total	87.28	65	218	283
% correct		83.08	88.53	

The random effects logit model was then fit where the two groups of branch effects were considered random. More specifically, we consider

$$\text{logit}\,(p_{ijklm}) = b_0 + u_{(i)} + BRN_{j(i)} + TY_k + NEW_l + b_1\,AMT + b_2\,AMT/TNW \qquad (5.8.2)$$

where $BRN_{j(1)} \sim N(0, \sigma_1^{\,2})$

$BRN_{j(2)} \sim N(0, \sigma_2^{\,2})$,

and the remaining variables are as defined in the fixed effects model.

These variances were estimated as

$$\sigma_1^{\,2} = 1.044$$
$$\sigma_2^{\,2} = 0.081$$

These values indicate a larger variability of the branch effect for the oversampled group than for the other 44 branches. However, this result (σ_1^2 = 1.044) may be questioned. First, in the oversampled group, we have a very small sample of 4 branches. These branches were not selected randomly but rather chosen by the division management of the bank. Thus, the results may not be representative of all the branches. Second, the loan applications retrieved from these 4 branches may not comprise the entire population of declines. There were no controls enforced on the sampling within the particular branches, and, therefore, a quantity of information may not have been made available for the analysis (see Section 3.3). These two factors would also distort the variance of the branch effect.

On the other hand, the value for σ_2^2 is a more reliable estimate. Again, this is the variance of the branch effect for the 44 branches which were sampled at the division office. The sampling procedure at this level allowed for greater controls over selection, and the sample itself was fairly large. The fact that the variance is small ($\sigma_2^2 = 0.0811$) indicates that there is little variability in the effect from branch to branch, and, consequently, the model provides a good fit to the data. If important variables had been omitted from the model, this would have increased the variance term.

To further illustrate the magnitude of σ_2^2, we can construct an interval for the probability of loan approval - the p's, (refer to Section 5.1.3). The interval covering plus and minus two standard deviations of the branch effect for p is

(0.7434 , 0.9005)

where the mean value for p is 0.8366.

Overall, there has been a large reduction in the variability (from 1.89 in Section 6.2 to 0.081 in this section). This indicates that much of the variability of the branch effect was due to the sampling procedure. In fact, the coefficient for the oversampling effect in the LRME model (Table A-12) is equal to -2.79. In terms of the value of p in the model, the odds of a loan approval, p divided by (1-p) - declines by 93.85 percent if the loan application is filed at one of the four branches in the branch sample of declines. This is due primarily to the fact that the sample of observations from these branches are comprised mostly of declines, relative to the remaining 44 branches.

The beta coefficients and the classification results for the LRME model are given in Table A-12 and Table A-11, respectively. In comparison to the estimates in Table A-4 for the original model, these estimates are relatively stable. When the sampling design was presented (in Section 3.5) it was noted that the branch effects were forced to be signigficant due to the oversampling. However, it was stressed that this result would have no bearing on the other variables in the model. It has now been confirmed that these other variables are consistent and unaffected by the sampling design.

One of the assumptions made in Chapter Four was that the branch effects come from a population which is normally distributed with a mean of zero. However, due to a small number of observations in some cells, it is possible that the branch effects estimates could be grouped around the mean. Thus, analyzing the distribution of the posterior estimates in an attempt to justify the normality assumption for the prior distribution is not a conclusive methodology. Nevertheless, we consider the estimates of the branch effects (under LRME) in the non-oversampled group (the division data) and study the shape of the distribution of these values. The reason is that this non-oversampled group is a representative sample from the population of branches and the cells in this group are similar in number of observations. The specific effects from the posterior distribution do not appear to be normally distributed. In fact, a

TABLE A-12

COEFFICIENTS FOR THE "SEPARATION" MODELS

		LRFE	LRFE*	LRME
Constant		1.874	4.1722	2.5507
TY	(1)	0	0	0
	(2)	-.477	.2435	-.0271
	(3)	.233	-.1724	.6654
	(4)	.644	.6434	1.1264
NEW	(1)	0	0	0
	(2)	-1.401	-1.4879	-1.1302
	(3)	1.293	1.3514	1.3586
	(4)	.283	.1917	.4443
AMT		-.00063	-.00148	-.000449
AMT/TNW		.00037	.00201	.000189
U	(1)	0	0	0
	(2)	-1.3983	-3.2704	-2.7949
$BRN_{j(1)}$	(2)	1.9216	2.1149	1.1893
	(4)	-1.0953	-1.1438	-.7292
	(11)	.1056	.0204	.1102
	(19)	0	0	-.5703
$BRN_{j(2)}$	(1)	-	-3.4584	.0106
	(5)	-	4.6740	-.0623
	(6)	-	-3.6844	.0143
	(10)	-	∞	.0298
	(12)	-	∞	.0013
	(14)	-	∞	.0092
	(15)	-	∞	.0492
	(17)	-	∞	.0066
	(18)	-	∞	.0463
	(21)	-	-3.7767	-.0125
	(22)	-	∞	.0179
	(24)	-	∞	.0759
	(26)	-	∞	.0318
	(27)	-	-4.5809	-.0391
	(28)	-	∞	.0192
	(30)	-	-6.3819	-.0622
	(31)	-	∞	.0385
	(32)	-	∞	.0039
	(33)	-	-6.0611	-.0655
	(34)	-	-6.0992	-.0469
	(35)	-	2 ∞	.0106
	(36)	-	4.0779	.0482
	(37)	-	-4.8452	-.0429
	(38)	-	-∞	-.0644
	(39)	-	∞	.0157
	(40)	-	-4.4195	-.0181
	(41)	-	∞	.0348
	(42)	-	-5.2986	-.0335
$BRN_{j(2)}$	(43)	-	∞	.0223
	(44)	-	∞	.0161
	(45)	-	∞	.0059
	(46)	-	-5.2935	-.0528
	(47)	-	-6.2931	-.0993
	(48)	-	∞	.0052

where U (1) = remaining 44 branches
(2) = oversampled 4 branches

$BRN_{j(1)}$ = specific branches within oversampled group

$BRN_{j(2)}$ = specific branches within remaining 44 branches group

LRFE* = is as defined in section 5.8

Histogram of Coefficients for Non-oversampled Group

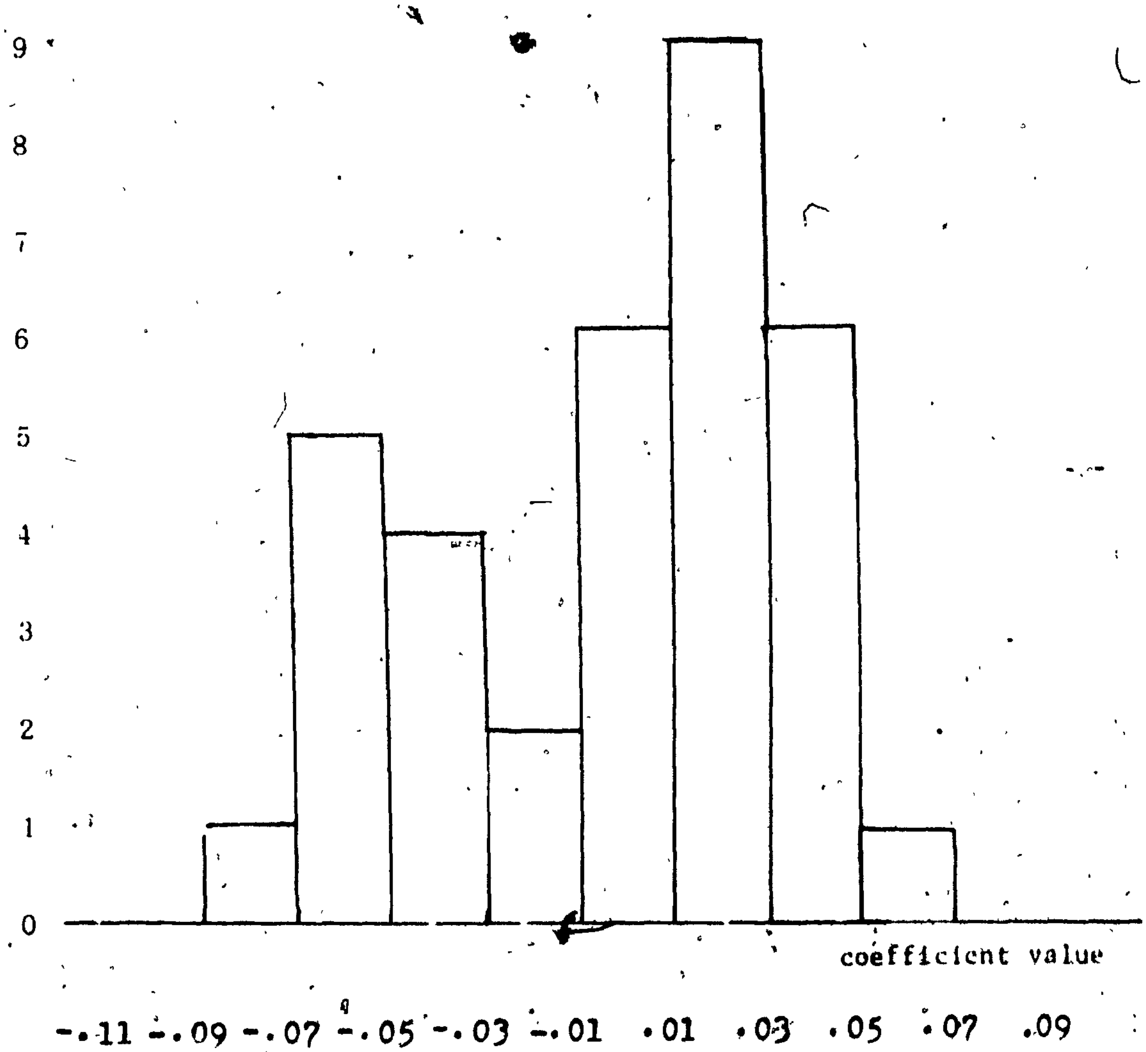


Figure 5.8.1

Lilliefors test for normality (Conover, 1980) was performed on these effects and a p-value of approximately 0.03 was obtained. The histogram for these BRN coefficients is provided in figure 5.8.1. It can be seen that the data seem to be bimodal rather than normal. However, this apparent departure from normality has not prevented the LRME from performing very well in prediction. This indicates that this model may well be robust with respect to normality departures and should be considered as a topic for further research.

In Table A-11, we see that, overall, the logit model with mixed effects performs with the same effectiveness as the fixed effects model (86.57 percent correct). Once again, the LRME does not do as well as the initial best model presented in Section 5.1.4 (87.99 percent compared to 86.57 percent correct), neither does the LRFE model do as well. The separation models in both cases did not fit the data as well as the original models (see Section 5.1). This may be due to a somewhat simplified model (applied LRME) where the BRN*TNW interaction term was removed from the model. (One must remember that these results relate to the classification of the data that was used in the model building. Prediction, classification of <u>new</u> data, is more difficult and results in a decline in effectiveness. There has been much evidence in related research which has indicated that this decline is greater in the fixed effects model than in the random effects model (see, for example, Efron and Morris (1975) ; this is illustrated in the holdout and bootstrap analysis in Sections 5.5 and 5.6.)

It should be noted that the separation models presented here to address the problem of oversampling, were limited to logistic regression where the application was rather straightforward. The same type of methodology could be extended to the discriminant model; however, this is left for future research.

The Likelihood Function

The likelihood function ($\ln p(y \mid X, \sigma^2)$) was investigated. In the separation model for LRME there are two prior variances, σ_1^2 and σ_2^2. The value of the likelihood function is maximized when σ_1^2 and σ_2^2 are set at the values determined by the EM algorithm (i.e., $\sigma_1^2 = 1.044$; $\sigma_2^2 = 0.0811$). As can be seen in Table A-13 (and in the chart reproduced below), the likelihood function is maximized at this point and falls off when moving in any direction away from this optimal point. (See "Points on Grid" where σ_1^2 is on the X axis, σ_2^2 is on the Y axis, and the values in the table represent the likelihood [$\ln p(y \mid X, \sigma^2)$] expressed as a percentage of its maximum.)

In addition, the plot illustrates the shape of the likelihood as a function of the prior variances. We see the likelihood values (Z - as a fraction of the likelihood at its maximum) drop off sharply as both σ_1^2 and σ_2^2 get smaller. As σ_2^2 gets larger and the first prior variance gets smaller, the decline is readily apparent but not as sharp as above. However, when σ_1^2 gets larger (> 1.044), the likelihood function flattens out (when σ_2^2 is small, large or stays at 0.0811). This indicates that the information in the data regarding σ_1^2 may not be as precise as it is for σ_2^2. This conclusion is consistent with the results discussed above.

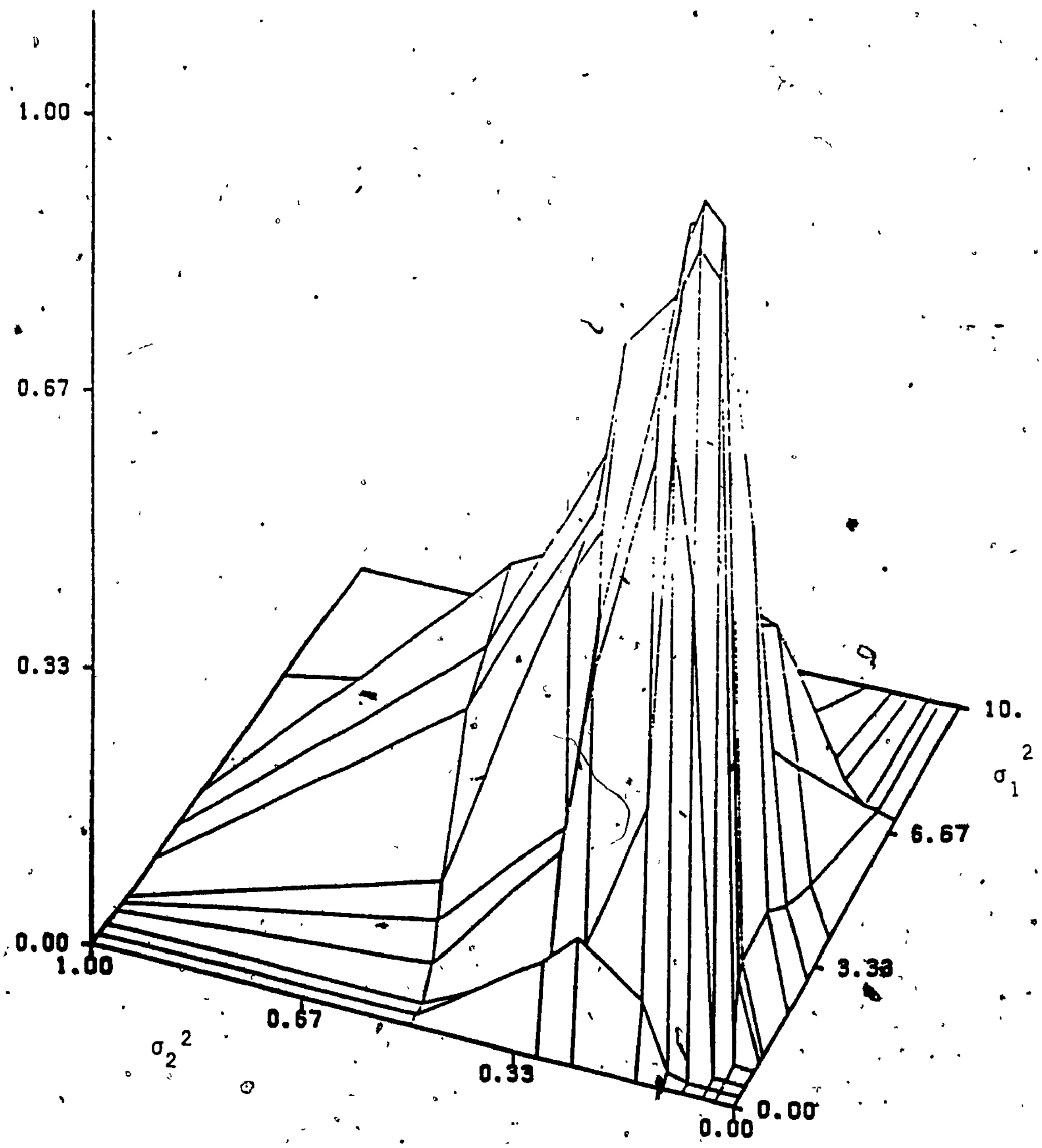
Table A-13

POINTS ON GRID

$\sigma_1^2(x)$ → / $\sigma_2^2(y)$	0.000	.030	.050	.0811	.110	.150	.250	.300	.500	1.000
0.000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.250	.0000	.0000	.0000	.0000	.0079	.0842	.1379	.1011	.0003	.0000
0.500	.0000	.0000	.00008	.5622	.7368	.2983	.1139	.0777	.0007	.0000
0.850	.0000	.0021	.9187	.9513	.8962	.6913	.4172	.2019	.0319	.0000
1.044	.0000	.0944	.9752	1.000	.9423	.7743	.4703	.2212	.0732	.0000
1.200	.0000	.1062	.8911	.9705	.9615	.8429	.5731	.5100	.1129	.0000
2.200	.0000	.1156	.5665	.5986	.6155	.8399	.7522	.5428	.2731	.0000
3.000	.0000	.0836	.2654	.4088	.4921	.6542	.6037	.5933	.3091	.0000
4.000	.0000	.0592	.1989	.3641	.3735	.3927	.4889	.4426	.3636	.0000
7.000	.0000	.0064	.0092	.0364	.0447	.0991	.1257	.0628	.0742	.0000
10.000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.00000

THREE-DIMENSIONAL PLOT OF LIKELIHOOD FUNCTION

Z - percentage of optimal point (1.044, 0.0811)



5.9 Multi-Model Approach

At this point, a number of possible methods in analyzing the bank data have been examined. In this section, we introduce another methodology in order to better fit the data.

In Chapter Three on the discussion of the raw data, we described two separate sampling designs. First, at the division office, 278 files were examined (252 were approved and 26 were declined) and decisions on these files were reached by division management. It should be remembered that all 278 cases were first recommended at the branch managers' decision level.

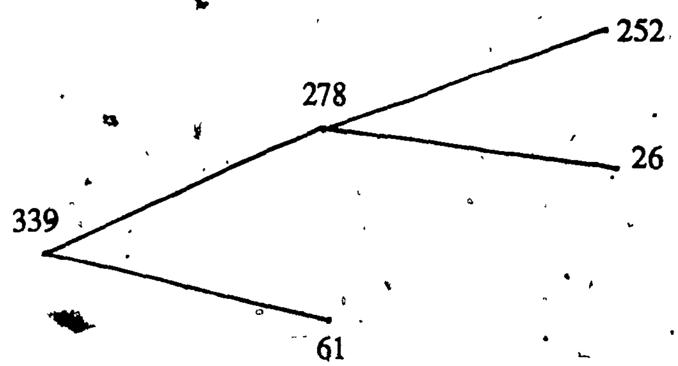


Figure 5.9.1

A second sample was selected of 61 cases (declines only) from the branch managers' offices from 4 pre-selected branches. These 61 declines, combined with the 278 previous cases (all recommended at the branch manager's level), provide information concerning the branch manager's decision activity.

As a result, in this section, we develop two models for the purpose of comparison and fit. First, we consider modelling the division manager's decision process by analyzing the first sample of 278 cases. Second, the branch manager's decision process is examined by considering all 339 cases; the second sample of 61 declines comprises all of the "no" group.

Division Management Level

A linear discriminant model (LDA) was fit to the original sample of 278 cases. The results from the stepwise analysis provided the following model:

$$Z_{ijk} = b_o + BRN_i + NEW_j + b_1 BRN_i{*}AMT \qquad (5.9.1)$$

where Z_{ijk} is the discriminant score which is compared to a predetermined cutoff point

BRN is the branch of the bank

NEW represents the experience of the applicant; and

BRN*AMT is the interaction of branch and the amount of the loan application.

A summary of the classification results, along with estimates of the coefficients, are given in Table A-14 and A-15. The model was able to correctly classify 92.45 percent of the cases.

Table A -14
Division Management Level Classification

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	30.77	8	18	26
	Yes	98.81	3	249	252
	Total	92.45	11	267	278
	% correct		72.73	93.26	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	58.33	14	10	24
	Yes	98.53	3	201	204
	Total	94.30	17	211	228
	% correct		82.35	95.26	
C.	Logit Regression (Mixed Effects) (LRME)				
	No	58.33	14	10	24
	Yes	91.18	18	186	204
	Total	87.72	32	196	228
	% correct		43.75	94.90	

TABLE A - 15

Coefficients for the Division Management Level

	LDA	LRFE	LRME
Constant	-.72134	8.1118	3.592
TY(1)		0	0
(2)		-.3783	-1.0481
(3)		-3.8687	.0175
(4)		.29724	-.1422
NEW(1)	0	0	0
(2)	.4560	-3.9002	-.9497
(3)	0	14.243.	-.0861
(4)	0	-1.3946	.0401
AMT		-.005481	-.00229
AMTTN		-.001405	.000122
BRN 1	.000	0	.2385
2	.000	.3651	.0775
4	-0.9712	.7821	.2300
5	.067	1.9840	.1498
6	.003	-2.3695	-.00079
10	-.092	2. ∞	.2238
11	1.5009	-3.6[illegible]	-.9814
12	.000	∞	.0235
14	.000	∞	.0695
15	.000	∞	.3885
17	.004	∞	.0795
18	.539	∞	.2502
19	.052	22.4	.1077
21	.000	-2.2102	-.0489
22	.007	∞	.1105
24	.000	∞	.6181
26	0.000	∞	.1259
27	0.6510	-.7830	.01101
28	-0.1700	∞	.1245
30	-2.8409	-6.3634	-.3071
31	-.001	∞	.4569
32	-.001	∞	.0243
33	.007	44.464	-.5166
34	9.2106	-19.902	-.3726
35	.051	∞	.1522
36	-.006	9.7753	.2979
37	2.3912	-34.377	-.4701
38	7.0192	-∞	-.6600
39	.071	∞	.0556
40	.000	-.02007	-.1053
41	.000	∞	.2229
42	-2.6863	58.902	-.2140
43	.005	∞	.1581
44	.007	∞	.1221
45	.001	∞	.0448

		LDA	LRFE	LRME
BRN	46	.000	363.91	-.2635
	47	.000	-2.2326	-.4685
	48	0	∞	.0451
BRN*TNW	1		0	-.0001
	2		.0005	.0004
	4		.0529	.0001
	5		.0417	.0001
	6		.0086	.0006
	10		.0349	.0001
	11		.0108	.0015
	12		.1009	.0000
	14		.0344	.0000
	15		-.0104	.0001
	17		-.5436	.0000
	18		-.00707	.0002
	19		-.1414	.0001
	21		.0049	.0008
	22		.0085	.0001
	24		.0097	.0005
	26		.0043	.0001
	27		-.0021	-.0002
	28		.00605	.0004
	30		.00807	-.0001
	31		.0060	.0003
	32		.0099	.0000
	33		-.7756	-.0002
	34		.2859	.0000
	35		-.0074	.0000
	36		-.0194	.0000
	37		.7561	.0000
	38		-1.4586	-.0001
	39		.0709	.0000
	40		.3911	.0002
	41		-.6971	.0001
	42		-1.364	-.0001
	43		.0470	.0000
	44		.1263	.0000
	45		.0246	.0000
	46		-3.4032	-.0002
	47		-.00491	-.0011
	48		-.3753	.0000
BRN*AMT	1	.00134		
	6	.00377		
	27	.00515		
	30	.01900		
	33	.01959		
	34	-.08766		
	42	.04260		
	46	.00851		
	47	.01705		

A logit model with fixed effects (LRFE) was run on the division management sample as well. The model used from this stepwise analysis is as follows:

$$\text{logit}\,(p_{ijkl}) = b_0 + BRN_i + TY_j + NEW_k + b_1\,AMT + b_2\,BRN_i{*}TNW \qquad (5.9.2)$$

where the variables are as defined in the earlier models. A list of classification results as well as parameter estimates is given in Tables A-14 and A-15. The overall classification results for this model show it has an ability to classify 94.30 percent of the cases correctly.

Following the same procedure employed in the earlier models, the best stepwise model for the fixed effects logit was also employed for the random effects. Thus we can present the LRME model as

$$\text{logit}(p_{ijkl}) = b_o + BRN_i + TY_j + NEW_k + b_1\,AMT + b_2\,BRN_i{*}TNW \qquad (5.9.3)$$

where $BRN_i \sim N(0,\sigma_1^2)$,

$BRN_i{*}TNW \sim N(0,\sigma_2^2)$

and the remaining variables are as defined above. Once again, a full listing of coefficients and classification results are given in the appendix, as well as in the text for convenience.

The variance estimates have decreased from the original models to σ_1^2 = 0.2143 and σ_2^2 = 0.00000273 in this model. This is due primarily to the fact that the oversampled cases were not being analyzed. Thus, the branch effect is fairly stable for the remaining cases. This is similar to what occurred in the "separation" models.

The classification results for the model show an effectiveness of 87.72 percent in classifying the same cases that were used in fitting the model. To put these classification figures in the right perspective, we can consider a naive model where all cases are classified as yes (the yes group is larger than the no group). If this was done, a naive model would classify 90.65 percent (252/278) correct. The table below summarizes the values for the decision management level and the results for the original model (Section 5.2).

	Division Management Percentage Correct	Original Model Percentage Correct
LDA	92.45	87.32
LRFE	94.30	91.87
LRME	87.72	87.99
Naive	90.65	74.34

It should be emphasized that these results pertain only to classification and not prediction. As has been documented in this thesis, prediction results may differ substantially and these results tend to favour the random or mixed effects model.

Branch Management Level

A stepwise linear discriminant analysis was performed on the branch management data. In this data set, there were only 4 branches which had declines; the remaining branches had all recommendations. All the declines were from the 61 cases in the second sample. All the observations in the first sample represented loan application recommendations at the branch manager's level.

Consequently, the BRN variable was expected to be highly significant in the variable selection. As a result, the best stepwise model was derived in two ways - the first method incorporated the BRN effect, and the other method ignored the effect of the branches. The first model is as follows:

$$Z_{ijk} = b_o + BRN_i + TY_j + NEW_k + TY_j{*}NEW_k + b_1\, BRN_i{*}AMT \qquad (5.9.4)$$

where Z_{ijk} is the discriminant score and all other variables are as defined previously.

The second model, which contains no branch effect, is as follows:

$$Z_{ijk} = b_o + TY_i + NEW_j + b_1\, AMT + b_2\, TNW \qquad (5.9.5)$$

The classification results and the listing of coefficients for both models are given in Tables A-16 and A-17.

The LDA was able to classify 87.61 percent correct with the "BRN effect included"; the model without the branch effect classified only 61.65 percent correct.

A logit model with fixed effects was also run on the branch manager's decision data. The stepwise procedure produced the following model:

$$\text{logit}\,(p_{ijkl}) = b_0 + BRN_i + TY_j + NEW_k + b_1\,AMT/TNW \qquad (5.9.6)$$

Following the same procedure as above, the BRN variable was removed from the stepwise selection process, and the effectiveness of the model was investigated. The resulting model is:

$$\text{logit}\,(p_{ijk}) = b_o + TY_i + NEW_j + b_1\,AMT/TNW + b_2\,TNW \qquad (5.9.7)$$

Once again, all coefficients and classification results are listed in Tables A-16 and A-17. The logit model incorporating the BRN effect was able to classify 94.67 percent correct, while the second logit model, with fixed effects could only classify 84.84 percent correct.

Table A-16

Branch Management Level Classification

	Observed	% correct	Classified No	Classified Yes	Total
A.	Linear Discriminant Analysis (LDA)				
	No	96.72	59	2	61
	Yes	85.61	40	238	278
	Total	87.61	99	240	339
	% correct		59.60	99.17	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	86.54	45	7	52
	Yes	96.88	6	186	192
	Total	94.67	51	193	244
	% correct		88.24	96.37	

Table A-16

Branch Management Level Without BRN.

	Observed	% correct	Classified No	Classified Yes	Classified Total
A.	Linear Discriminant Analysis (LDA)				
	No	83.33	45	9	54
	Yes	56.44	98	127	225
	Total	61.65	143	136	279
	% correct		31.47	93.38	
B.	Logit Regression (Fixed Effects) (LRFE)				
	No	51.92	27	25	52
	Yes	93.75	12	180	192
	Total	84.84	39	205	244
	% correct		69.23	87.80	

TABLE A - 17

Coefficients for Branch Management Level

	LDA	LRFE
Constant	1.77919	-11.414
TY 1	-.16611	0
2	0	-10.089
3	0	.30915
4	0	-1.7746
NEW 1	-.29352	0
2	0	2.5643
3	0	.70215
4	0	-.22422
AMTTN		-.03300
BRN 1	.20974	0
2	-1.07855	8.2058
4	-3.4707	13.195
5	-.44011	-1.1929
6	-.22332	.52068
10	-1.0058	∞
11	-2.8849	10.443
12	-.92781	∞
14	-.37665	∞
15	.1024	∞
17	-.62869	∞
18	-.37811	∞
19	-3.9314	13.322
21	-.21344	.58144
22	1.0424	∞
24	-.41573	∞
26	.11139	∞
27	-.60398	2.2406
28	1.04236	∞
30	-.37255	2.2728
31	-.45154	∞
32	-.45079	∞
33	.23582	.08561
34	-.51350	9.3727
35	-.8300	∞
36	-.30728	-.74049
37	.47308	-.58427
39	.09621	∞
40	1.04236	.88700
41	1.04236	∞
42	-.50651	2.4276
43	-.35701	∞
44	-.54956	∞
45	-.73368	∞
46	-.58385	1.1872

		LDA	LRFE
BRN	47	-.66766	9.3801
	48	0	∞
TYNEW	(1,1)	0	
	(1,2)	-.35811	
	(1,3)	0	
	(1,4)	-.03510	
	(2,1)	-.00228	
	(2,2)	-1.70853	
	(2,3)	-.61775	
	(2,4)	-.42875	
	(3,1)	1.07208	
	(3,2)	.28735	
	(3,3)	-.6169	
	(3,4)	-.02315	
	(4,1)	-.09545	
	(4,2)	-.10092	
	(4,3)	-1.87501	
	(4,4)	0	
BRN*AMT	1	-.00056	
	2	.00062	
	3	-.04479	
	4	.00017	
	5	.00426	
	6	.00051	
	7	0	
	8	-.00828	
	9	-.02289	
	10	.00519	
	11	.00219	
	12	.03744	

Coefficients for Branch Management Level without BRN

		LDA	LRFE
Constant		1.43587	0.24961
TY.	1	-2.00149	0
	2	0	-10.100
	3	0	-1.9412
	4	0	-2.5206
NEW	1	-.29461	0
	2	-1.3095	.0[illegible]796
	3	0	-.88481
	4	0	-2.0277
TNW		.00032	-.00107
AMTTN		—	-.00446
AMT		.00045	—

In order to evaluate these results, we can once again apply a type of "challenging" or "sophisticated" naive model based on the BRN variable. If we classify as a no all the cases coming from the four oversampled branches and place all other cases in the yes group, we arrive at the following results:

Model	Using BRN Percentage Correct	Not Using BRN Percentage Correct
LDA	87.61	61.65
LOGIT	94.67	84.84
Naive	82.07	74.34

The naive model without branch is simply the method of classifying all cases in the yes group (the modal group).

A logit model with random effects was not incorporated due to the nature of the data set. Recall only four branches had declines, and this does not constitute a large enough sample to warrant random effects.

Comparison of Results

In this section, a number of different models have been presented. Although the variables which were found to be significant were not always the same as those identified in the original analysis (see Section 5.2), they were, for the most part, highly consistent with the earlier results. At both the division management and the branch management levels, the logit model with fixed effects performed better than its competitors. The same success was exhibited by the original models that dealt with classification. In some instances a naive model was shown to have performed well.

Classification results are given below for the Division level, the Branch level, and the Overall level decision process (original model - Section 5.1).

Percentage Correct

	Original	Division	Branch	Percentage Combined
LDA	87.32	92.45	87.61	82.540
LRFE	91.87	94.30	94.67	91.169
LRME	87.99	87.72	-----	-----
Naive	74.34	90.65	82.07	74,620

The models (except for LRME) appear to do better when the data are dissected and analyzed separately by two decision processes. For example, for LRFE, the overall effectiveness of the original model was 91.87 percent. However, when the two decision processes are analyzed separately, the effectiveness increases - 94.30 and 94.67 percent.

One explanation for this behaviour may involve the sampling scheme of the data. The 61 cases were drawn in a non-random fashion and represent declines at a different point in the decision process at the bank. The original models used the information in those cases to retrieve information on the method of arriving at a decline at any level of the decision process. If it is believed that the decision processes differ at the two levels, then a two-level model may be more appropriate. However, the premise in this research is that the decision processes are similar, and, hence, all the data were pooled before being analyzed (Section 5.1).

In order to arrive at a "bottom line" comparison, we would like to know what percentage of cases would be classified correctly in this data set. One method for this evaluation is the following: Since there are 82 percent recommendations and 18 percent rejections at the branch level, the "branch classification effectiveness" is multiplied by weights of 0.82 and 0.18. Once a loan has been recommended, it must be submitted to the division management for approval where the level of effectiveness (for yes and no decisions separately) is multiplied by the number of applications received. We can use the LRFE model to demonstrate this comparative technique. At the branch level, 96.88 percent of the yes and 86.54 percent of the no applications are evaluated correctly which is comprised of approximately 18% rejections and 82% approvals. The ones which are approved must then be classified at the division level where the model's effectiveness is 98.53 percent for the yes's and 58.33 percent correct for the no's. The combination of these stages results in an overall classification of 91.169 percent correct. This level is slightly less than the original in classification effectiveness (91.87 percent). (Again, emphasis should be placed on the word classification - these are not prediction results.)

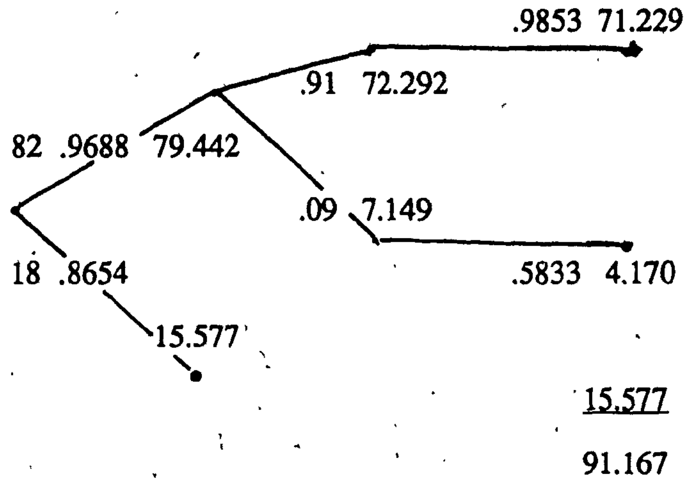


Figure 5.9.2

This section illustrates a second alternative method to the modelling of the bank data. In some respects, the different approaches are very competitive when compared to the original model (Section 5.1). However, the original model does perform well in classification and prediction. True comparisons should be made only after cross validation techniques and residual analysis have been applied to all the models; but this study will be reserved for future research.

It is hoped that these techniques will continue to be applied in credit scoring research and other relevant areas in an effort to develop better modelling strategies.

CHAPTER SIX

CONCLUSIONS

6.1 Summary

The results presented in Chapter Five give a very detailed description of the effectiveness of the three different models presented in this paper across a number of criteria. The discussion in this chapter concentrates on which method is best overall.

The logit regression with mixed effects model (the LRME) performance is superior to its competitors with respect to prediction of new cases in the holdout and bootstrap samples. It also showed superiority in the residual analysis.

One important factor to be considered in the comparison of the LRME and LDA models is the assumptions inherent in multivariate normal linear discriminant analysis (see Sections 2.7 and 4.1). Even though some of these assumptions were not justified in this credit scoring application, the LDA model has proven to be quite robust. These findings are in agreement with those presented by Reichert et al (1983).

With repect to the LRME, we were able to produce a useful model. In Section 5.8, we presented a random effects model which takes into consideration the sampling design of this study. In future research, loan applications could be scored using this model; the $u_{(i)}$ term or the "oversampling effect" could be dropped. (Presumably this effect will not be evident in other sample designs.)

6.2 Limitations of this Research

In Chapter One, it was clearly stated that the primary objective of this research was to introduce a new methodology relevant to the analysis of the credit scoring problem. This new methodology took the form of applying a state of the art statistical method to a well-known financial problem. We then compared the effectiveness of this methodology with previous methods employed. Although we cannot state that logit regression with mixed effects is the best method for all credit scoring research, we have shown that, in this particular application, it is quite effective and is superior to its competitors.

It is highly possible that this random effects model will perform better in future credit scoring applications. Further, it is highly probable that the exact model determined here (with respect to significant variables and values of the parameters and coefficients) will not be repeated in other studies. However, this does not diminish the importance of this research. A new method has been introduced to handle the task of analyzing the important variables in the decision process of the commercial loan officer. This method was proven to be quite fruitful and should be investigated further.

6.3 Suggestions for Further Research

Further research into the area of logit models and random effects models should be pursued. New variable selection techniques (more cost-efficient) should be investigated to determine the best models for LRME. At present, only an approximation from the fixed effects model was used.

Second, larger and more detailed data sets should be obtained and analyzed. Perhaps more conclusive results could be reached if more data were available. It has been discussed that subjective data (Section 2.4) should be included in the analysis. This information could be obtained through personal interviews with the bank (or credit) managers.

Discriminant analysis with random effects should also be considered. The random effects definitely improved the effectiveness of the logit model. It is possible that this would do the same for LDA resulting in an overall superior model.

Finally, research could be extended into other areas of credit scoring. Recalling from Chapter Two, we limited our study to commercial loan applicants and to the decision of the loan officer to approve or decline the loan. The methodology set forth in this paper could be extended to commercial credit outside the banking industry. As well, bankruptcies and defaults in loans are topics which deserve consideration. The area of consumer credit is another possible application of this research.

In short, this dissertation has established a framework for future research into the development of more effective and efficient credit scoring models.

BIBLIOGRAPHY

Altman, E.I.(1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", The Journal of Finance, September.

Altman, E.I.(1973), "Predicting Railroad Bankruptcies in America", Bell Journal of Economics and Management Science, Spring.

Altman, E.I., Haldeman, R.,and Narayanan, P.(1977), "Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations", Journal of Banking and Finance, June.

Altman, E.I., and Eisenbeis, R.A.(1978),"Financial Applications of Discriminant Analysis", Journal of Financial and Quantitative Analysis, March.

Ameniya, T. (1981), "Qualitative Response Models: A Survey", Journal of Economic Literature, December.

Bierman, H., and Hausman, W.H.(1970), "The Credit Granting Decision", Management Science, Vol 16, April.

Biomedical Computing Programs (BMDP-83) P-Series (1983), Editors W.J. Dixon and M.B. Brown, University of California Press.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975), Discrete Multivariate Analysis: Theory and Practice, The MIT Press.

Blum, M.(1974),"Failing Company Discriminant Analysis", Journal of Accounting Research, Spring.

Chang, P.C., and Afifi, A.A. (1974),"Classification Based on Dichotomous and Continuous Variables", Journal of the American Statistical Association, June.

Cohen, K.J., Gilmore, T.C., and Singer, F.A. (1966),"Bank Procedures for Analyzing Business Loan Applications", in Analytical Methods in Banking, eds. K.J. Cohen and F.S. Hammer, R.D. Irwin, Inc.

Conover, W.J. (1980), Practical Nonparametric Statistics, Second Edition, John Wiley and Sons.

Cox, D.R. (1970), Analysis of Binary Data, Meuthen and Co. Limited.

Dawes, R.M., and Corrigan, B. (1974),"Linear Models in Decision Making", Pschological Bulletin, p. 95-106.

Deakin, E.B.(1972),"A Discriminant Analysis of Predictors of Business Failure", Journal of Accounting Research, Spring.

Degroot, M.H. (1970), Optimal Statistical Decisions, McGraw Hill Book Company.

Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977),"Maximum Likelihood from Incomplete Data Via the EM Algorithm", Journal of the Royal Statistical Society, Volume B 39.

Dempster, A.P., Rubin, D.B., and Tsutakawa, R.K. (1981), "Estimation in Covariance Components Models", Journal of the American Statistical Association, June.

Doreen, D.D., and Farhoomand, F.(1983),"A Decision Model for Small Business Loans", Journal of Small Business (Canada), Fall.

Durand, D. (1941), "Risk Elements in Consumer Installment Financing", National Bureau of Economic Research, Study #8.

Edmister, R.O.(1971),"Financial Ratios and Credit Scoring for Small Business Loans", The Journal of Commercial Bank Lending, September.

Edmister, R.O.(1972),"An Empirical Test of Financial Ratio Analysis fro Small Business Failure Prediction", Journal of Financial and Quantitative Analysis, March.

Efron, B.(1982),"The Jackknife, the Bootstrap and Other Resampling Plans", Society for Industrial and Applied Mathematics, 38.

Efron, B. (1983),"Estimating the Error Rate of a Prediction Rule: Improvement of Cross-Validation", Journal of the American Statistical Association, June.

Efron, B. and Morris, C. (1972), "Limiting the Risk of Bayes and Empirical Bayes Estimators - Part II: The Empirical Bayes Case", Journal of the American Statistical Association, March.

Efron, B. and Morris, C. (1973), "Stein's Estimation Rule and Its Competitors - An Empirical Bayes Approach", Journal of the American Statistical Association, March.

Efron, B. and Morris, C. (1975), "Data Analysis Using Stein's Estimator and Its Generalizations", Journal of the American Statistical Association, June.

Eisenbeis, R.A.(1977),"Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics", The Journal of Finance, June.

Eisenbeis, R.A., and Avery, R.B. (1972),"Discriminant Analysis and Classification Procedures", Lexington Books.

Fienberg, S.E. (1980), The Analysis of Cross-Classisfied Categorical Data, Second Edition, The MIT Press.

Gong, G. (1986),"Cross-Validation, the Jackknife and the Bootstrap: Excess Error Estimation in Forward Logistic Regression", Journal of the American Statistical Association, March.

Goodman, L.A. (1978), Analyzing Qualitative/Categorical Data, Edited by J. Magidson, Abt Books.

Green, P.E. (1978), Analyzing Multivariate Data, The Dryden Press.

Haberman, S.J. (1978), Analysis of Qualitative Data, Academic Press, Volume I and II.

Hatch, J., Wynant, L., and Grant, M.J. (1985),"Bank Lending under the Small Business Loans Act", Canadian Banker, October.

International Mathematical and Statistics Library (1982), Edition #9, IMSL, Inc., Houston, Texas, June revision.

James, W., and Stein, C. (1961),"Estimation with Quadratic Loss", Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1, University of California Press.

Joy, O.M., and Tollefson, J.O.(1975),"On the Financial Application of Discriminant Analysis", Journal of Financial and Quantitative Analysis, December.

Joy, O.M., and Tollefson, J.O.(1978),"Some Clarifying Comments on Discriminant Analysis", Journal of Financial and Quantitative Analysis, March.

Laird, N.M. (1975),"Log-linear Models with Random Parameters: An Empirical Bayes Approach", Ph.D. Dissertation, Harvard University.

Laird, N.M. (1978), "Empirical Bayes Methods for Two-Way Contingency Tables", Biometrika, Vol. 65.

Landwehr, J.M., Pregibon, D., and Shoemaker, A.C. (1984), "Graphical Methods for Assessing Logistic Regression Models", Journal of the American Statistical Association, March.

Leonard T. (1975), "Bayesian Estimation Methods for Two-Way Contingency Tables", Journal of the Royal Statistical Society, Volume B 37.

McFadden, D. (1976),"A Comment on Discriminant Analysis 'versus' Logit Analysis", Annals of Economic and Social Measurement, volume 5.

Neter, J., Wasserman, W., and M.H. Kutner (1985),Applied Linear Statistical Models, Second Edition, Richard D. Irwin, Inc.

Orgler, Y.E.(1970),"A Credit Scoring Model for Commercial Loans", Journal of Money, Credit and Banking, November.

Orgler, Y.E.(1971),"Evaluation of Bank Consumer Loans with Credit Scoring Models", Journal of Bank Research, Spring.

Reichert, A.K., Cho, C.C., and Wagner, G.M.(1983),"An Examination of the Conceptual Issues Involved in Developing Credit-Scoring Models", Journal of Business and Economic Statistics, April.

Rosenbaum, P.R., and Rubin, D.B., (1983) "The Central Role of the Propensity Score in Observational Studies", Biometrika, Vol 70.

Rubin, D.B. (1984),"Graphical Methods for Assessing Logistic Regression Models: A Comment", Journal of the American Statistical Association, March.

SAS Institute Inc. (1985), SAS User's Guide: Graphics, Version 5 Edition, Cary, N.C., SAS Institute Inc.

Sprague, R.H., and Watson, H.J. (1976), "A Decision Support System for Banks", Omega - The International Journal of Management Science, Vol. 4, No. 6.

Szonyi, A.J., and Steinhoff, D. (1979),Small Buiness Management Fundamentals, 1st Canadian Edition, McGraw-Hill Ryerson Limited.

Tomberlin, T.J. (1982),"A Statistical Perspective on Predicting Losses in Automobile Insurance", Ph.D. Dissertation, Harvard University.

Tomberlin, T.J. (1987), "Predicting Accident Frequencies for Drivers Classified by Two Factors", Journal of the American Statistical Association, to appear.

Van Horne, J.C. (1977), Financial Management and Policy, Fourth Edition, Prenctice Hall, Inc.

Wiginton, J.C.(1980),"A Note on the Comparison of Logit and Discriminant Models of Consumer Credit Behaviour", Journal of Financial and Quantitative Analysis, September.

Wong, G.Y., and Mason, W.M. (1985),"The Hierarchical Logistic Regression Model for Multilevel Analysis", Journal of the American Statistical Association, September.

THE QUALITY OF THIS MICROFICHE IS HEAVILY DEPENDENT UPON THE QUALITY OF THE THESIS SUBMITTED FOR MICROFILMING.

UNFORTUNATELY THE COLOURED ILLUSTRATIONS OF THIS THESIS CAN ONLY YIELD DIFFERENT TONES OF GREY.

LA QUALITE DE CETTE MICROFICHE DEPEND GRANDEMENT DE LA QUALITE DE LA THESE SOUMISE AU MICROFILMAGE.

MALHEUREUSEMENT, LES DIFFERENTES ILLUSTRATIONS EN COULEURS DE CETTE THESE NE PEUVENT DONNER QUE DES TEINTES DE GRIS.

APPENDIX A

TABLE A-1

CORRELATION MATRIX

		ACC 1	OUT 2	AMT 3	CRE 4	NEW 5	BRN 6	CLA 7	INCO 8	MO 9	TY 10	SE 11	ST 12	SA 13
ACC	1	1.000												
OUT	2	-0.258	1.000											
AMT	3	-0.249	-0.187	1.000										
CRE	4	-0.273	-0 115	0.867	1 000									
NEW	5	-0.099	-0.013	0 056	0 146	1.000								
BRN	6	0.246	0.084	-0 433	-0.458	-0.171	1.000							
CLA	7	0.116	0.083	-0 244	-0 236	-0.095	0.314	1.000						
INCO	8	0.040	0.069	-0.116	-0 110	-0 144	0.182	0.133	1.000					
MO	9	-0 199	0 130	0 003	0.056	0 047	-0 039	-0 120	0.025	1.000				
TY	10	0.020	-0.010	-0.051	-0.152	-0.301	0.209	0.201	0.132	0.048	1.000			
SE	11	0 000	0 000	0 000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
ST	12	-0.038	-0.059	0.175	0.180	0.410	-0.318	-0.170	-0 162	-0 238	-0.296	0 000	1.000	
SA	13	-0.132	-0 031	0 406	0 336	0 220	-0.207	-0 009	-0 067	0 037	-0 124	0 000	0.121	1 000
GP	14	-0.071	-0.017	0.231	0 190	0.192	-0.071	0.040	-0.023	0 089	-0.126	0 000	0.069	0 841
INC	15	-0 114	-0 035	0.263	0.146	0 087	-0 078	0 031	0 029	0 088	-0.085	0 000	-0.055	0 739
BD	16	-0 114	-0 029	0 164	0 147	0 007	-0 118	-0.009	0 326	0 117	0 089	0 000	-0.075	0.180
DEP	17	0 050	-0 031	0 204	0 156	0 089	-0 007	0 044	-0.009	0.133	0 055	0 000	-0 036	0 622
TRA	18	-0.102	-0.060	0 238	0 126	0.093	-0.089	0.042	0.033	0.059	-0 102	0 000	-0 033	0 652
TNW	19	-0.187	0 094	0 430	0 471	0 133	-0 224	-0 231	-0 087	0 051	-0.207	0 000	0 142	0 345
WC	20	-0.150	0 113	0 376	0 540	0 065	-0 250	-0.142	-0 047	-0.004	-0.196	0 000	0 053	0 036
TF	21	-0 002	-0.099	0 062	-0.045	-0.079	-0 107	-0 155	-0 071	-0 136	0.052	0.000	-0 073	-0.119
NO	22	-0.041	-0.012	0 056	0 169	0.016	-0 074	-0.117	0.015	0.108	0.016	0.000	0 047	-0 043
SEC	23	-0 039	0.043	0.281	0.247	-0.004	-0 165	-0 123	-0.046	0 008	-0.180	0.000	0.054	-0.001
AGE	24	0.032	-0.063	0.078	0 083	-0 022	-0.111	-0.135	-0.109	0 031	-0.111	0.000	0 120	0 073
FS	25	0 021	0 028	-0 049	-0 050	0 065	0 077	0.069	-0 015	-0 028	0.022	0 000	-0.122	-0.036
PUR	26	-0.033	-0.048	-0.065	-0.140	-0.286	0.194	0.117	-0.036	0.062	0.559	0.000	-0.292	-0.185

		GP 14	INC 15	BD 16	DEP 17	TRA 18	TNW 19	WC 20	TF 21	NO 22	SEC 23	AGE 24	FS 25	PUR 26
GP	14	1.000												
INC	15	0.811	1.000											
BD	16	0.114	0.223	1.000										
DEP	17	0.812	0.611	0 095	1.000									
TRA	18	0.698	0.950	0.198	0.464	1.000								
TNW	19	0.309	0.284	0.027	0 163	0.298	1 000							
WC	20	-0.120	-0.099	0.022	-0.236	-0.025	0.735	1.000						
TF	21	-0.256	-0 123	-0.208	-0.198	-0.076	-0 072	-0.010	1.000					
NO	22	-0.014	-0 021	0 271	0 025	-0.014	0.066	0 057	-0.294	1.000				
SEC	23	0.041	0.025	-0.034	0.018	0.032	0 130	0.204	0.015	-0.065	1.000			
AGE	24	0.064	0.025	-0.095	0.150	0 006	0 070	0.026	-0.022	0 045	0.009	1.000		
FS	25	-0.023	-0 021	-0 015	-0 031	-0 013	-0 037	-0 020	0 021	0.054	-0.014	-0.133	1.000	
PUR	26	-0.121	-0.044	-0.084	0.000	-0 069	-0 119	-0.113	-0 003	-0 041	-0 149	-0.040	0 036	1.000

TABLE A-2 PRINCIPAL COMPONENTS ANALYSIS

SORTED ROTATED FACTOR LOADINGS (PATTERN)

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	FACTOR 9
GP	14	0.940	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
INC	15	0.928	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SA	13	0.856	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TRA	18	0.845	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DEP	17	0.800	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
BRN	6	0.000	0.724	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AMT	3	0.260	-0.663	0.527	0.000	0.000	0.000	0.000	0.000	0.000
CRE	4	0.000	-0.626	0.609	0.000	0.000	0.000	0.000	0.000	0.000
CLA	7	0.000	0.560	0.000	0.000	0.000	0.000	0.000	-0.310	0.000
WC	20	0.000	0.000	0.870	0.000	0.000	0.000	0.000	0.000	0.000
TNW	19	0.252	0.000	0.779	0.000	0.000	0.000	0.000	0.000	0.000
PUR	26	0.000	0.000	0.000	0.823	0.000	0.000	0.000	0.000	0.000
TY	10	0.000	0.000	0.000	0.818	0.000	0.000	0.000	0.000	0.000
NO	22	0.000	0.000	0.000	0.000	0.762	0.000	0.000	0.000	0.000
TF	21	0.000	-0.250	0.000	0.000	-0.734	0.000	0.000	0.000	0.000
INCO	8	0.000	0.000	0.000	0.000	0.000	0.771	0.000	0.000	0.000
BD	16	0.000	-0.252	0.000	0.000	0.335	0.712	0.000	0.000	0.000
OUT	2	0.000	0.309	0.000	0.000	0.000	0.000	0.726	0.000	0.000
ACC	1	0.000	0.417	0.000	0.000	0.000	0.000	-0.710	0.000	0.000
MO	9	0.000	0.000	0.000	0.000	0.000	0.000	0.356	0.668	0.000
ST	12	0.000	-0.307	0.000	-0.397	0.000	0.000	0.000	-0.618	0.000
FS	25	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.814
AGE	24	0.000	0.000	0.000	0.000	0.000	-0.305	0.000	0.000	-0.613
SEC	23	0.000	0.000	0.378	-0.309	0.000	0.000	0.000	0.316	0.000
SE	11	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NEW	5	0.000	0.000	0.000	-0.452	0.000	0.000	0.000	-0.336	0.000
	VP	4.149	2.417	2.287	1.983	1.505	1.474	1.401	1.280	1.195

THE ABOVE FACTOR LOADING MATRIX HAS BEEN REARRANGED SO THAT THE COLUMNS APPEAR IN DECREASING ORDER OF VARIANCE EXPLAINED BY FACTORS. THE ROWS HAVE BEEN REARRANGED SO THAT FOR EACH SUCCESSIVE FACTOR, LOADINGS GREATER THAN .5000 APPEAR FIRST. LOADINGS LESS THAN .2500 HAVE BEEN REPLACED BY ZERO.

TABLE A-3

TABLE A-3.

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	F VALUE TO ENTER OR REMOVE	NUMBER OF VARIABLES INCLUDED	U-STATISTIC	APPROXIMATE F-STATISTIC	DEGREES OF FREEDOM	
1	49 D26		71.0186	1	.8259	71.019	1.00	337.00
2	42 D19		42.6572	2	.7329	61.227	2.00	336.00
3	57 D34		35.3900	3	.6629	56.793	3.00	335.00
4	92 D107		24.2174	4	.6181	51.601	4.00	334.00
5	135 D148		7.4316	5	.6046	43.562	5.00	333.00
6	149 D162		7.3129	6	.5915	38.209	6.00	332.00
7	129 D142		6.1929	7	.5807	34.147	7.00	331.00
8	144 D157		5.3275	8	.5714	30.935	8.00	330.00
9	132 D145		5.3556	9	.5623	28.456	9.00	329.00
10	140 D153		5.0230	10	.5538	26.426	10.00	328.00
11	72 D49		4.5192	11	.5463	24.692	11.00	327.00
12	136 D149		4.6801	12	.5385	23.279	12.00	326.00
13	35 D9		4.1945	13	.5317	22.022	13.00	325.00
14	148 D161		3.6271	14	.5258	20.873	14.00	324.00
15	28 D2		3.2290	15	.5206	19.831	15.00	323.00
16	93 D108		3.0000	16	.5158	18.894	16.00	322.00

BINARY VARIABLES DEFINED

```
IF (NEW EQ 1) THEN D1=1
IF (NEW EQ 2) THEN D2=1
IF (NEW EQ 3) THEN D3=1
IF (TY EQ 1 ) THEN D4=1
1PAGE    4   BMDP7M

IF (TY EQ 2) THEN D5=1
IF (TY EQ 3) THEN D6=1
IF (TY EQ 5) THEN D8=1
IF (TY EQ 6) THEN D9=1
IF (PUR EQ 1) THEN D10=1
IF (PUR EQ 2) THEN D11=1
IF (PUR EQ 3) THEN D12=1
IF (BRN EQ 1) THEN D16=1
IF(BRN EQ 1) THEN D116=AMT
IF (BRN EQ 2) THEN D17=1
IF (BRN EQ 2) THEN D117=AMT
IF (BRN EQ 3) THEN D18=1  IF (BRN EQ 3) THEN D118=AMT
IF (BRN EQ 4) THEN D19=1  IF (BRN EQ 4)  THEN D119=AMT
IF (BRN EQ 5) THEN D20=1  IF (BRN EQ 5) THEN D120=AMT
IF (BRN EQ 6) THEN D21=1  IF (BRN EQ 6) THEN D121=AMT.
IF (BRN EQ 7) THEN  D22=1  IF (BRN EQ 7) THEN D122=AMT.
IF (BRN EQ 8) THEN D23=1. IF (BRN EQ 8) THEN D123=AMT
IF (BRN EQ 9) THEN D24=1. IF (BRN EQ 9) THEN D124=AMT
IF (BRN EQ 10) THEN D25=1  IF (BRN EQ 10) THEN D125=AMT
IF (BRN EQ 11) THEN D26=1  IF (BRN EQ 11) THEN D126=AMT
IF (BRN EQ 12) THEN D27=1  IF (BRN EQ 12) THEN D127=AMT
IF (BRN EQ 13) THEN D28=1  IF (BRN EQ 13) THEN D128=AMT
IF (BRN EQ 14) THEN D29=1  IF (BRN EQ 14) THEN D129=AMT
IF (BRN EQ 15) THEN D30=1  IF (BRN EQ 15) THEN D130=AMT
IF (BRN EQ 16) THEN D31=1  IF (BRN EQ 16) THEN D131=AMT
IF (BRN EQ 17) THEN D32=1  IF (BRN EQ 17) THEN D132=AMT
IF (BRN EQ 18) THEN D33=1  IF (BRN EQ 18) THEN D133=AMT
IF (BRN EQ 19) THEN D34=1  IF (BRN EQ 19) THEN D134=AMT
IF (BRN EQ 20) THEN D35=1  IF (BRN EQ 20) THEN D135=AMT
IF (BRN EQ 21) THEN D36=1  IF (BRN EQ 21) THEN D136=AMT
IF (BRN EQ 22) THEN D37=1  IF (BRN  EQ 22) THEN D137=AMT
IF (BRN EQ 23) THEN D38=1  IF (BRN EQ 23) THEN D138=AMT
IF (BRN EQ 24) THEN D39=1  IF (BRN EQ 24) THEN D139=AMT.
IF (BRN EQ 25) THEN D40=1  IF (BRN EQ 25) THEN D140=AMT
IF (BRN EQ 26) THEN D41=1  IF (BRN EQ 26) THEN D141=AMT.
IF (BRN EQ 27) THEN D42=1  IF (BRN EQ 27) THEN D142=AMT
IF (BRN EQ 28) THEN D43=1  IF (BRN EQ 28) THEN D143=AMT
IF (BRN EQ 29) THEN D44=1  IF (BRN EQ 29) THEN D144=AMT
IF (BRN EQ 30) THEN D45=1  IF (BRN EQ 30) THEN D145=AMT
IF (BRN EQ 31) THEN D46=1  IF (BRN EQ 31) THEN D146=AMT.
IF (BRN EQ 32) THEN D47=1  IF (BRN EQ 32) THEN D147=AMT
IF (BRN EQ 33) THEN D48=1. IF (BRN EQ 33) THEN D148=AMT.
IF (BRN EQ 34) THEN D49=1  IF (BRN EQ 34) THEN D149=AMT
IF (BRN EQ 35) THEN D50=1  IF (BRN EQ 35) THEN D150=AMT
IF (BRN EQ 36) THEN D51=1  IF (BRN EQ 36) THEN D151=AMT
IF (BRN EQ 37) THEN D52=1  IF (BRN EQ 37) THEN D152=AMT.
IF (BRN EQ 38) THEN D53=1  IF (BRN EQ 38) THEN D153=AMT
IF (BRN EQ 39) THEN D54=1  IF (BRN EQ 39) THEN D154=AMT

IF (BRN EQ 40) THEN D55=1  IF (BRN EQ 40) THEN D155=AMT.
IF (BRN EQ 41) THEN D56=1  IF (BRN EQ 41) THEN D156=AMT
IF (BRN EQ 42) THEN D57=1  IF (BRN EQ 42) THEN D157=AMT.
IF (BRN EQ 43) THEN D58=1  IF (BRN EQ 43) THEN D158=AMT
IF (BRN EQ 44) THEN D59=1. IF (BRN EQ 44) THEN D159=AMT.
IF (BRN EQ 45) THEN D60=1. IF (BRN EQ 45) THEN D160=AMT.
IF (BRN EQ 46) THEN D61=1  IF (BRN EQ 46) THEN D161=AMT
IF (BRN EQ 47) THEN D62=1  IF (BRN EQ 47) THEN D162=AMT
D101=0  D102=0.  D103=0  D104=0. D105=0
1PAGE    5   BMDP7M

D106=0. D107=0  D108=0. D109=0. D110=0
D111=0. D112=0. D113=0  D114=0  D115=0
IF (TY EQ 1 AND NEW EQ 1) THEN D101=1
IF (TY EQ 2 AND NEW EQ 1) THEN D102=1.
IF (TY EQ 3 AND NEW EQ 1) THEN D103=1.
IF (TY EQ 4 AND NEW EQ 1) THEN D104=1
IF (TY EQ 2 AND NEW EQ 2) THEN D105=1
IF (TY EQ 1 AND NEW EQ 2) THEN D106=1.
IF (TY EQ 3 AND NEW EQ 2) THEN D107=1
IF (TY EQ 4 AND NEW EQ 2) THEN D108=1.
IF (TY EQ 1 AND NEW EQ 3) THEN D109=1.
IF (TY EQ 2 AND NEW EQ 3) THEN D110=1
IF (TY EQ 3 AND NEW EQ 3) THEN D111=1.
IF (TY EQ 4 AND NEW EQ 3) THEN D112=1
IF (TY EQ 1 AND NEW EQ 4) THEN D113=1
IF (TY EQ 2 AND NEW EQ 4) THEN D114=1.
IF (TY EQ 3 AND NEW EQ 4) THEN D115=1.
```

TABLE A-4

		Best Model Coefficients: LDA*	Best Model Coefficients: LDA	Best Model Coefficients: LRFE	Best Model Coefficients: LRME	Missing Data Models: LRFE	Missing Data Models: LRME
CONSTANT		-2.3152	-1.25716	4.9269	1.9548	4.8441	2.2070
TY	(1)	.5394	0	0	0	0	0
	(2)	1.3097	-.93754	1.2091	.4381	1.1254	.3738
	(3)	2.7312	0	-.3274	.8109	-.5972	.3040
	(4)	0	0	1.4260	1.2590	1.3519	.9798
NEW	(1)	.4932	0	0	0	0	0
	(2)	1.2285	.5505	-2.3281	-1.0363	-1.9930	-1.0269
	(3)	.55154	0	2.9120	1.185	1.9614	1.2336
	(4)	0	0	.7524	.5617	.64681	.51925
BRN	(1)	.1765	0	0	.8210	0	.6055
	(2)	.9667	0	-1.2619	-.2556	-1.1670	-.1949
	(3)	1.3874	0	--	.0056	--	.0704
	(4)	3.0836	2.5004	-5.4090	-2.5364	-4.9235	-2.1782
	(5)	.9186	0	3.5395	.5996	3.7791	.4764
	(6)	-.1272	0	-2.0780	.32192	-2.1015	.22128
	(7)	.0000	0	--	.0007	--	.0233
	(8)	.0000	0	--	.0783	--	.0989
	(9)	.0000	0	--	.0453	--	.0007
	(10)	.5055	0	--	.6800	--	.554
	(11)	2.8073	-2.34511	-5.0604	-2.345	-5.3829	-2.5259
	(12)	1.1686	0	--	.0423	--	.4329
	(13)	.0000	0	--	.2307	--	.2556
	(14)	.7855	0	--	.2584	--	.3578
	(15)	.1557	0	--	1.0276	--	.8447
	(16)	.0000	0	--	.0895	--	.1019
	(17)	1.1227	0	--	.1995	--	.3051
	(18)	1.0978	0	--	.8902	--	.7522
	(19)	4.1249	2.9247	-5.4316	-2.322	-5.0593	-2.3384
	(20)	.0000	0	--	.1202	--	.1023
	(21)	.6371	0	-1.6921	.18347	-1.3838	.3589
	(22)	.0000	0	--	.5268	--	.4171
	(23)	.0000	0	--	.1108	--	.09000
	(24)	.3586	0	--	1.1471	--	1.0351
	(25)	.0000	0	--	.0565	--	.0035

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
BRN (26)	-.4832	0	∞	.6665	∞	.7156
(27)	.1456	0	-1.7335	.0504	-1.4387	.1348
(28)	.0000	0	∞	.4449	∞	.3748
(29)	.0000	0	∞	.0763	∞	.0029
(30)	-.9044	0	-1.3677	-.6744	-2.2225	-.5680
(31)	.4851	0	∞	.8318	∞	.7940
(32)	-1.6012	0	∞	.1279	∞	.3726
(33)	.4845	0	48.497	-.8153	48.422	-.7168
(34)	6.6881	6.11798	-19.881	-.6022	-20.049	-.6539
(35)	1.3296	0	∞	.2321	∞	.2083
(36)	1.1094	0	7.8243	.7798	-7.4096	.1434
(37)	-.6767	0	-11.265	-.2718	-3.8630	-.0740
(38)	.0000	0	-∞	-.9312	-∞	-.9428
(39)	.0000	0	∞	.4076	∞	.3181
(40)	.5553	0	-1.1681	.0474	-.96126	.2179
(41)	.6309	0	∞	.6816	∞	.6896
(42)	-1.32383	0	59.468	-.2756	49.458	.3040
(43)	.4325	0	∞	.5147	∞	.4720
(44)	1.4828	0	∞	.3098	∞	.3525
(45)	.6240	0	∞	.1875	∞	.1544
(46)	-1.1373	0	213.99	-.4274	199.56	-.4166
(47)	1.9594	0	-2.7595	-.6926	-2.8093	-.8304
(48)	0	0	∞	.1694	∞	.4336
BRN*TNW (1)			0	.000033	0	.000042
(2)			.000469	.000505	.000618	.00044
(3)			–	.000032	–	.00001
(5)			.03731	.00016	.03448	.0001123
(6)			.00821	.000989	.00777	.000775
(10)			.04895	.000152	.0447	.000108
(7)			–	.000451	–	.00011
(8)			–	.001007	–	.00009
(9)			–	.000098	–	.000775
(11)			.0140	.00349	.0157	.00325
(12)			.0995	.00000204	.07174	.000053
(14)			.0629	.0000495	.0271	.000045

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
BRN*TNW (13)			—	.006801	—	- .00015
(15)			- .00801	.00022	- .00760	.000166
(17)			- .16021	.0000063	- .00994	.000041
(16)			—	.000367	—	.00072
(18)			- .006898	.000206	- .006127	.00016
(19)			.00005738	.0000264	.0002518	.000048
(21)			.00322	.000756	.00296	.00010
(20)			—	.00008	—	.00001
(22)			.02047	.000347	.0162	.000244
(24)			.00567	.000453	.0059	.000354
(23)				.00075	—	.00083
(26)			.00432	.000251	.00444	.000375
(25)			—	.000277	—	.000036
(27)			- .000338	.0001177	- .0003711	- .00000957
(28)			.0113	.000836	.00752	.00062
(30)			- .0219	- .0002815	- .0106	- .00019
(29)			—	.0043	—	.0044
(31)			+ .00388	.00010	.00434	.000309
(32)			.00949	.000083	- .0116	.000087
(33)			- .81095	- .000227	- .80668	- .0001844
(34)			.28359	.0000799	.28394	.0000504
(35)			.00194	.000055	- .00250	.000039
(36)			- .0151	.000039	1.3944	.00010
(37)			.213	.00023	.0548	.000156
(38)			- .6487	- .0000143	- .967	- .00001286
(39)			.09165	.0000547	.0918	.0000378
(40)			.3958	.000394	.38677	.000332
(41)			- .00355	.000169	- .00323	.000119
(42)			-1.3828	- .0000778	-1.165	- .000081
(43)			.00675	.000043	.01009	.000038
(44)			.17018	.0000163	.15057	.000015
(45)			.00138	.0000251	0	.000018
(46)			-1.9887	- .000182	- 1.8532	- .00015
(47)			- .00424	- .0011177	- .00424	- .00100
(48)			- .15747	.0000106	.0203	.000075
(4)			- .00518	- .000448	.004518	.00069

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
NEW*TY (1,1)	-.4967	0				
(1,2)	-11.2596	0				
(1,3)	-2.3042	0				
(1,4)	.0000	0				
(2,1)	.0000	0.66337				
(2,2)	.0000	-0.67239				
(2,3)	.2396	0				
(2,4)	-2.1615	0				
(3,1)	-3.4704	0				
(3,2)	.0000	0				
(3,3)	.0000	0				
(3,4)	.0000	0				
(4,1)	-1.6621	0				
(4,2)	-2.1082	0				
(4,3)	-2.4964	0				
(4,4)	0	0				
AMT			- .003572	- .001497	- .003791	- .001381
AMT/TNW			.0013773	.001559	.006566	.002583
BRN*AMT (1)	.0010	0				
(2)	.0000	0				
(3)	.0000	0				
(4)	-.0001	0				
(5)	-.0051	0				
(6)	.0032	0				
(7)	.0255	0				
(8)	.0134	0				
(9)	.0294	0				
(10)	-.0026	0				
(11)	-.0011	[illegible]				
(12)	-.0376	[illegible]				
(13)	.0120	[illegible]				
(14)	.0144	[illegible]				
(15)	-.0002	[illegible]				
(16)	-.04[illegible]	0				
(17)	-.0141	0				
(18)	-.0035	0				

	BEST MODEL COEFFICIENTS				MISSING DATA MODELS	
	LDA*	LDA	LRFE	LRME	LRFE	LRME
BRN*AMT (19)	-.0024	0				
(20)	.0060	0				
(21)	-.0007	0				
(22)	-.0022	0				
(23)	-.1734	0				
(24)	-.0008	0				
(25)	-.0510	0				
(26)	.0023	0				
(27)	.0038	.00375				
(28)	-.0015	0				
(29)	.0201	0				
(30)	.0113	.00611				
(31)	-.0004	0				
(32)	-.0734	0				
(33)	.0126	.0122				
(34)	-.0539	-.05374				
(35)	-.0385	0				
(36)	-.0011	0				
(37)	.0159	0				
(38)	.2474	.22925				
(39)	-.0347	0				
(40)	.00156	0	$\sigma_1^2 = 1.8928$			$\sigma_1^2 = 1.6979$
(41)	.0000	0				
(42)	.0301	.01592				
(43)	.0007	0				
(44)	-.0247	0				
(45)	.0000	0	$\sigma_2^2 = .000004164$			$\sigma_2^2 = .00000331$
(46)	.0124	.00506				
(47)	.0059	.01224				
(48)	0	0				

NOTE:

TY (1) - operating loan
(2) - loan for equipment
(3) - loan for land and/or building
(4) - some combination of all three

NEW (1) - new firm with bank history
(2) - new firm with no bank history
(3) - existing firm with bank history
(4) - exisitng firm with no bank history

TABLE A-5

Summary Table

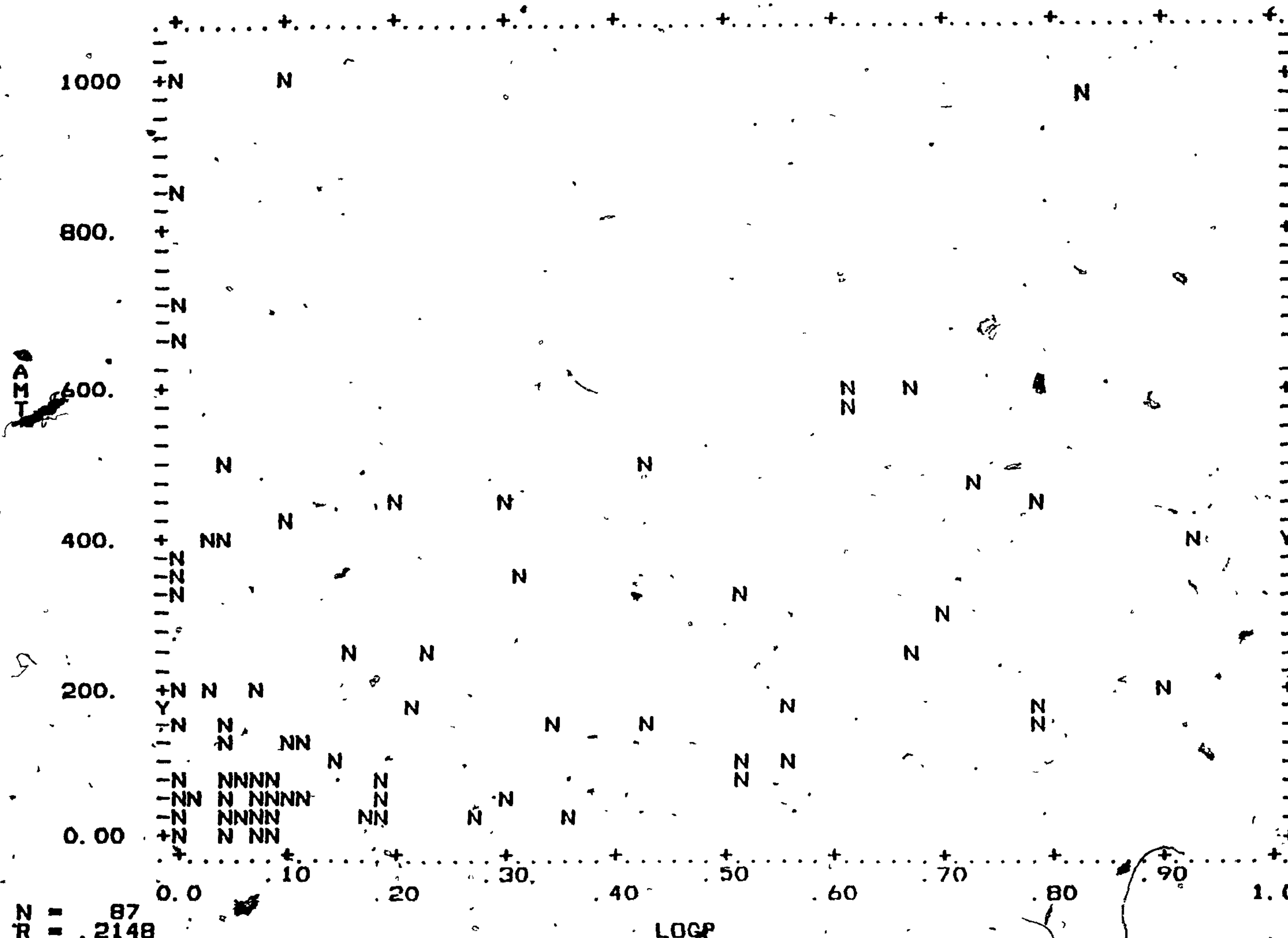
TERM	APPROX. CHI-SQ. ENTER	D.F.	APPROX. CHI-SQ. REMOVE	D.F.	P-VALUE	LOG LIKELIHOOD
TY			16.91	6	.0096	-61.7590
TY			IS IN			MAY NOT BE REMOVED.
NEW			31.16	3	.0000	-69.8854
NEW			IS IN			MAY NOT BE REMOVED.
ERN			86.81	37	.0000	-96.7124
ERN			IS IN			MAY NOT BE REMOVED.
AMT			7.25	1	.0071	-56.9329
AMT			IS IN			MAY NOT BE REMOVED.
TNN	.10	1			.7570	-53.2581
TNN	IS OUT					MAY NOT BE ENTERED.
C*D	15.46	37			9993	-45.5758
C*D	IS OUT					MAY NOT BE ENTERED.
C*E			61.65	37	.0067	-81.1289
C*E			IS IN			MAY NOT BE REMOVED.
AMTTN			2.30	1	.1292	-54.4568
AMTTN			IS IN			MAY NOT BE REMOVED.
CONSTANT			25.21	1	.0000	-65.9111
CONSTANT			IS IN			MAY NOT BE REMOVED.

NO TERM PASSES THE REMOVE AND ENTER LIMITS (.1500 .1000)

TABLE A-6.1

TABLE A-6.1.1

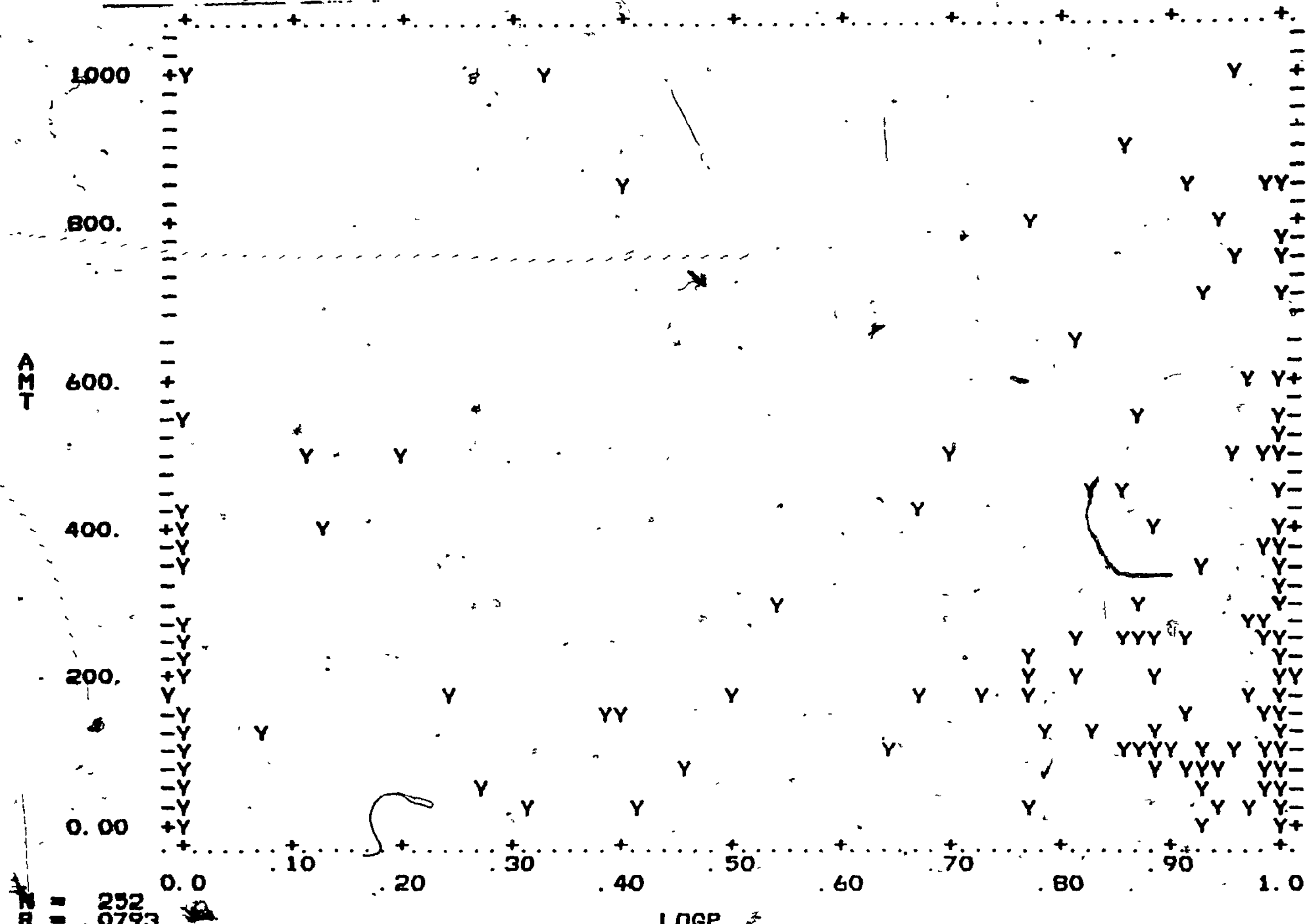
LRFE



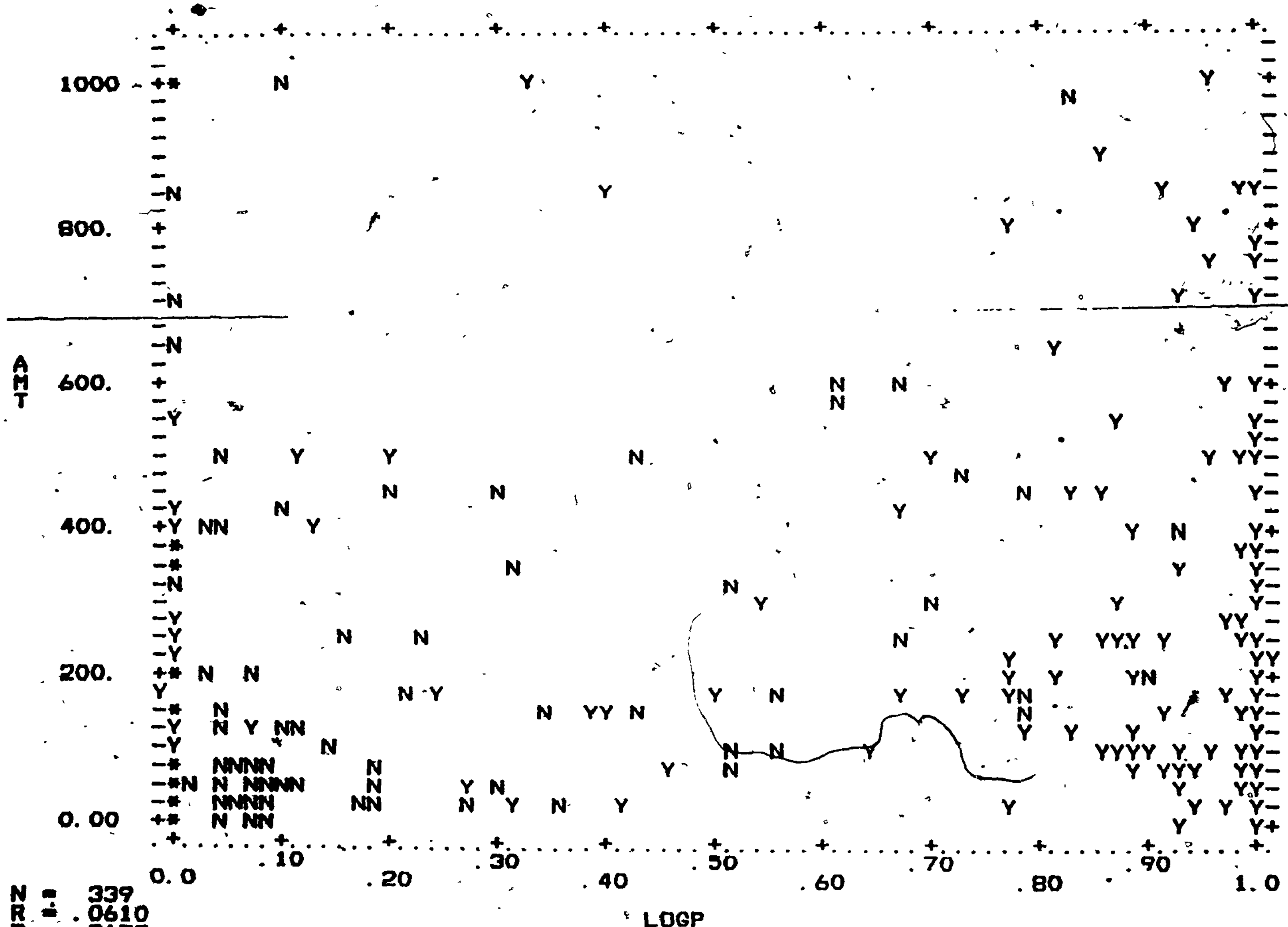
N = 87
R = .2148
P = .0455

--REGRESSION LINE-- --RMS-- MEAN S.D.
Y=-179.36 +210.27*X 63510. X .21469 .26203
Y 224.51 256.53

LOGP VERSUS AMT (28 VS. 3) GROUP=ND , SYMBOL=N
1PAGE 7 BMDPAD PLOTS ANAL OF BDATA




```
N =   252
R = .0793
P = .2101
                                           MEAN     S.D.
--REGRESSION LINE--     --RMS--       X   .73554   .39681
Y= 164.72 +43.713*X      47792.       Y   196.88   218.87

LOOP       VERSUS AMT         ( 28 VS.  3 )  GROUP=YES     , SYMBOL=Y
1PAGE    8  BMDP6D PLOTS ANAL OF BDATA
```

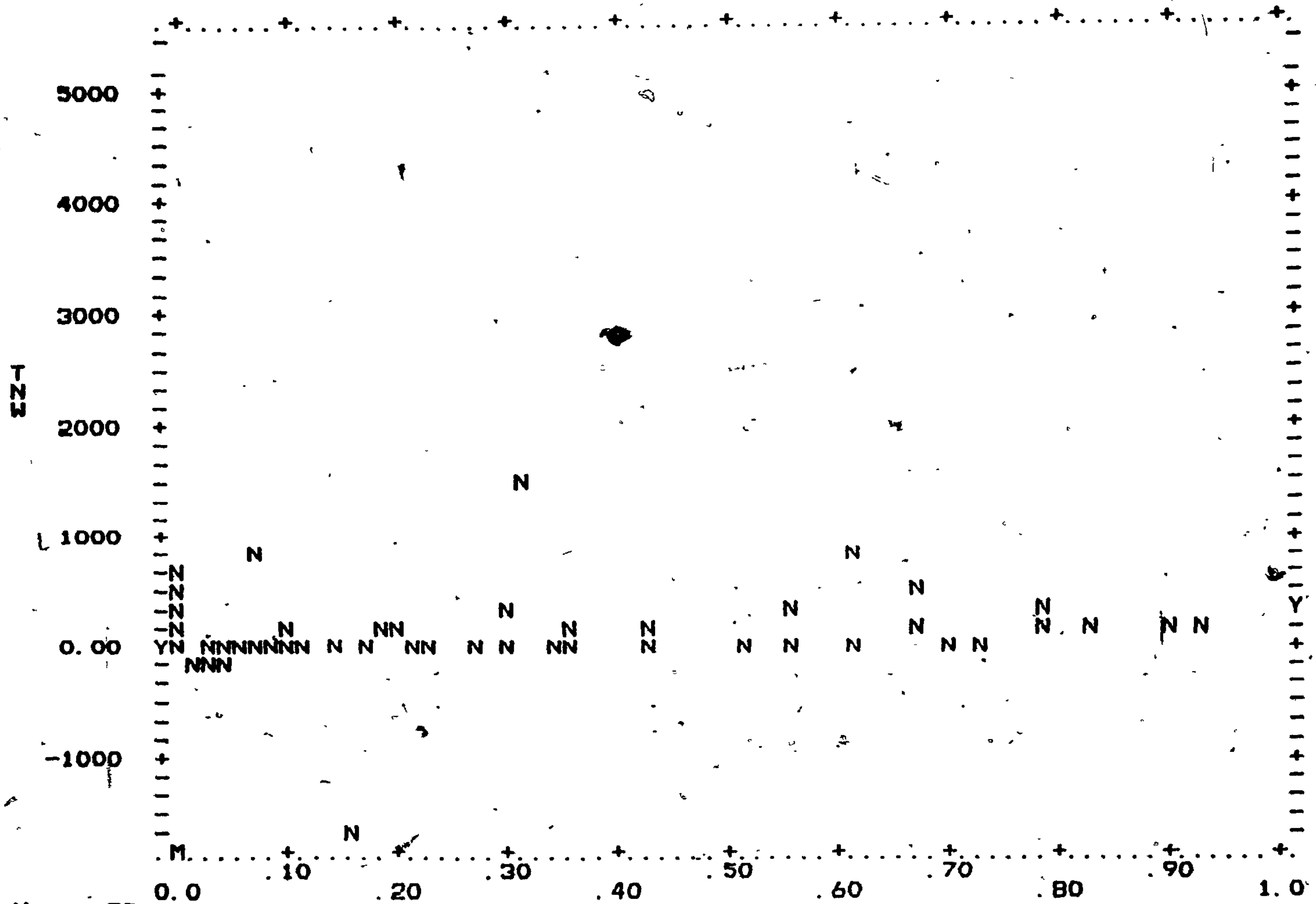


N = 339
R = .0610
P = .2629

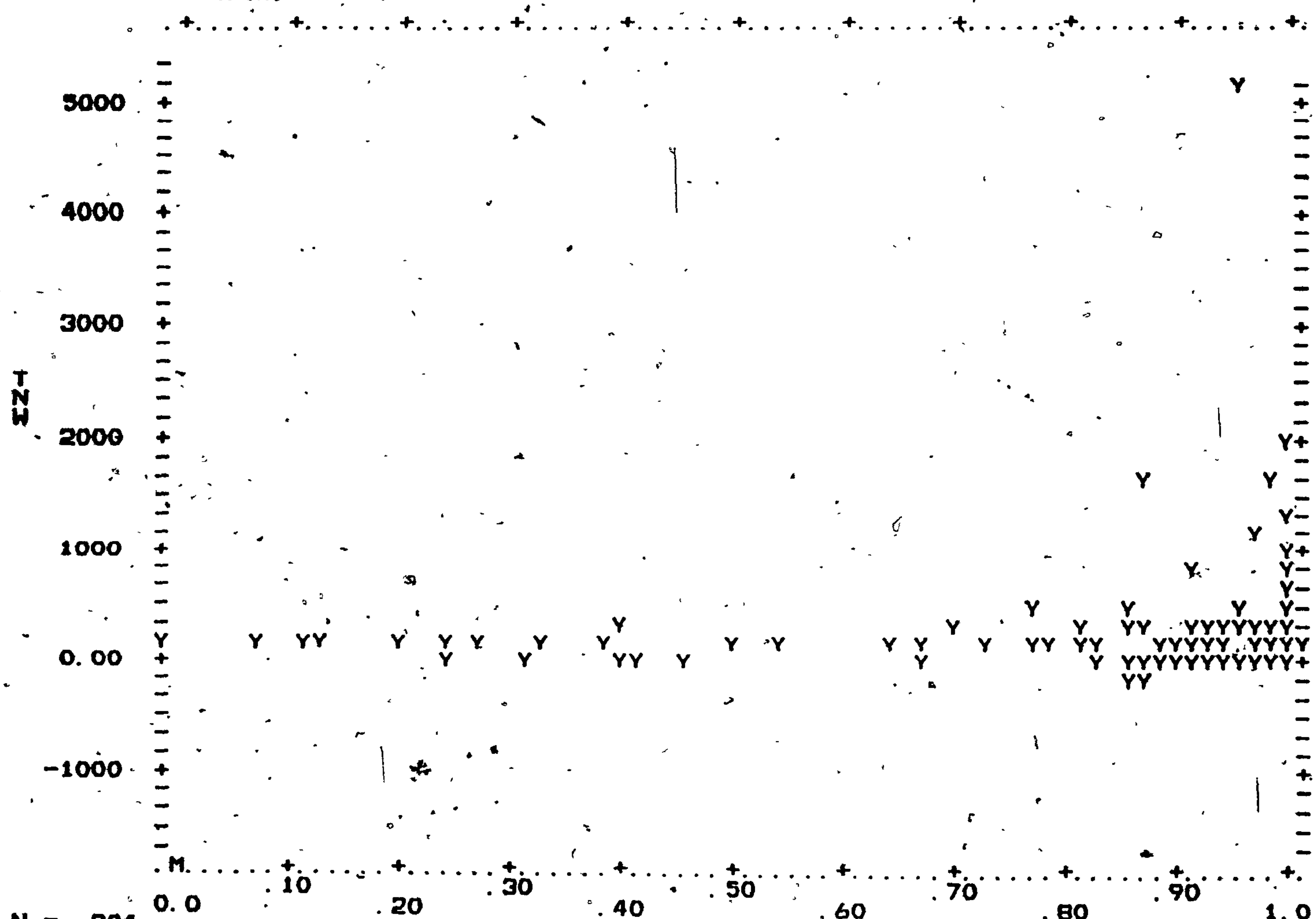
--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y= 184.48 +32.372*X	52422.	X	.60187	.43163
		Y	203.97	229.05

LOGP VERSUS AMT (28 VS. 3) GROUP=NO , SYMBOL=N
GROUP=YES , SYMBOL=Y

1PAGE 9 BMDP6D PLOTS ANAL OF BDATA



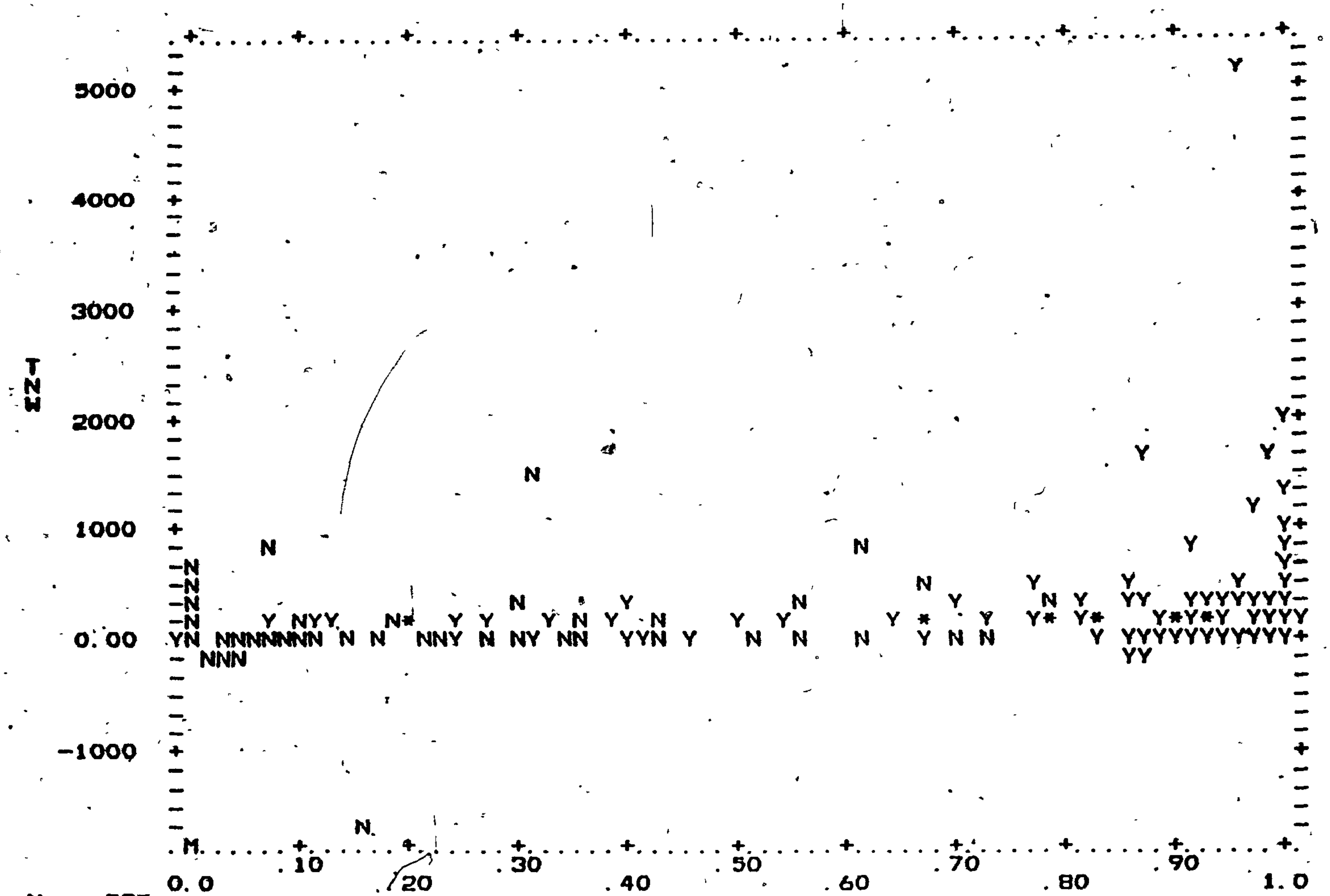

```
N =    79
R =  .1921
P =  .0900
                                              MEAN      S.D.
--REGRESSION LINE--     --RMS--      X     .23643    .26551
Y= 43.333 +226.67*X      95781.      Y     96.924    313.33

LOGP     VERSUS TNW        ( 28 VS. 19 )  GROUP=NO        , SYMBOL=N
1PAGE   19  BMDP6D PLOTS ANAL OF BDATA
```



N = 204
R = .0427
P = .5449

--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y= 113.39 +101.00*X	206009	X	.90861	.19148
		Y	205.16	453.18

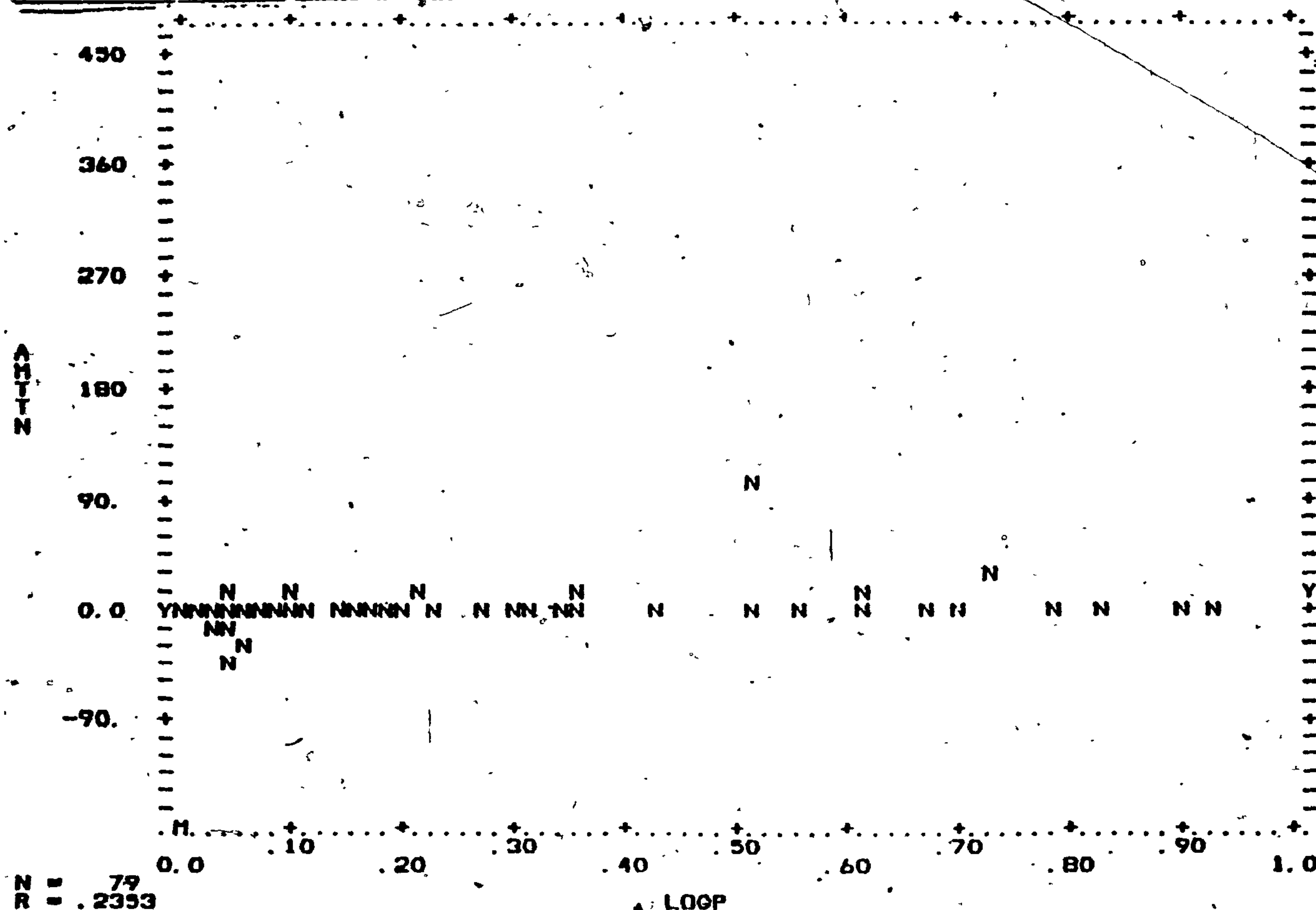
LOGP VERSUS TNW (28 VS. 19) GROUP=YES SYMBOL=Y
1PAGE 20 BMDP6D PLOTS ANAL OF BDATA



N = 283
R = .1396
P = .0187

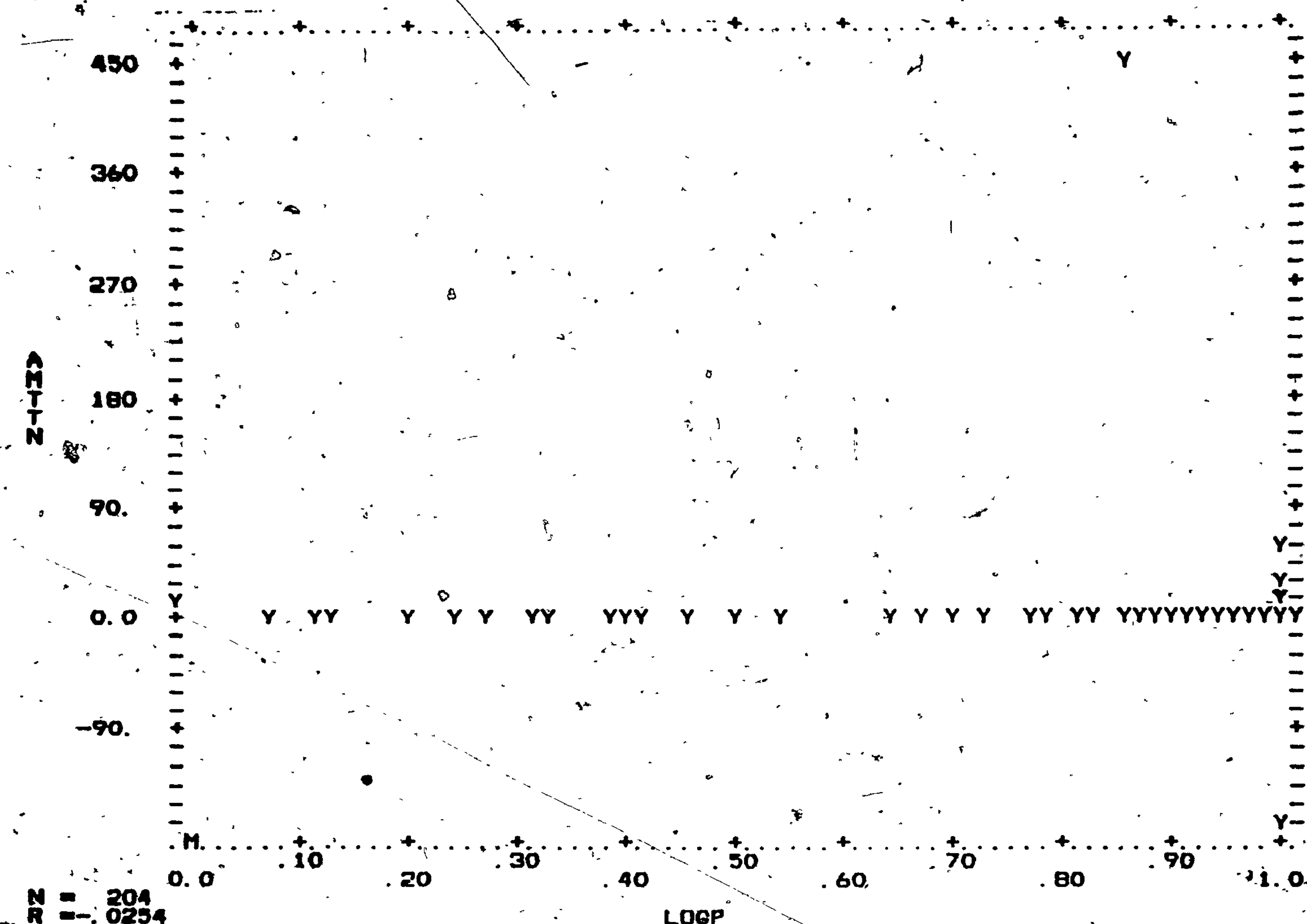
--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y= 60.453 +158.81*X	174517	X	.72097	.37032
		Y	174.95	421.14

LOGP VERSUS TNW (28 VS. 19) GROUP=NO , SYMBOL=N
GROUP=YES , SYMBOL=Y



N = 79
R = .2353
P = .0366

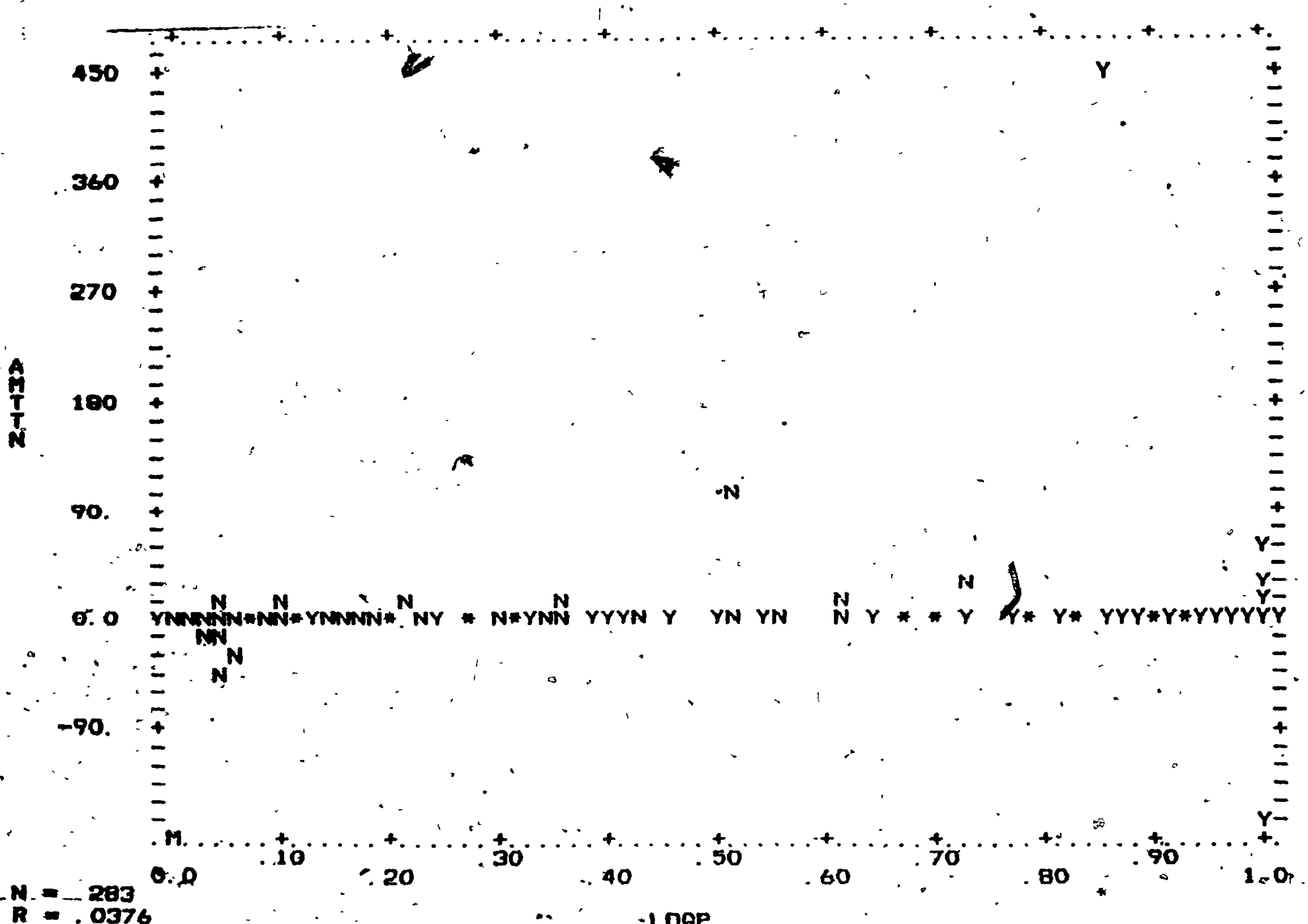
```
--REGRESSION LINE--      --RMS--        MEAN      S.D.
Y=-1.3100 +13.081*X      208.44    X   .23643    .26551
                                   Y  1.7827    14.759

LOOP     VERSUS AMTTN      ( 28 VS. 31 )   GROUP=NO      SYMBOL=N
1PAGE  25  BMDP6D PLOTS ANAL OF BDATA
```



N = 204
R =-.0254
P = .7189

```
--REGRESSION LINE--      --RMS--        MEAN      S.D.
Y= 7.9231 -4.4983*X       1156.4     X   .90861    .19148
                                     Y  3.8359   33.933
```

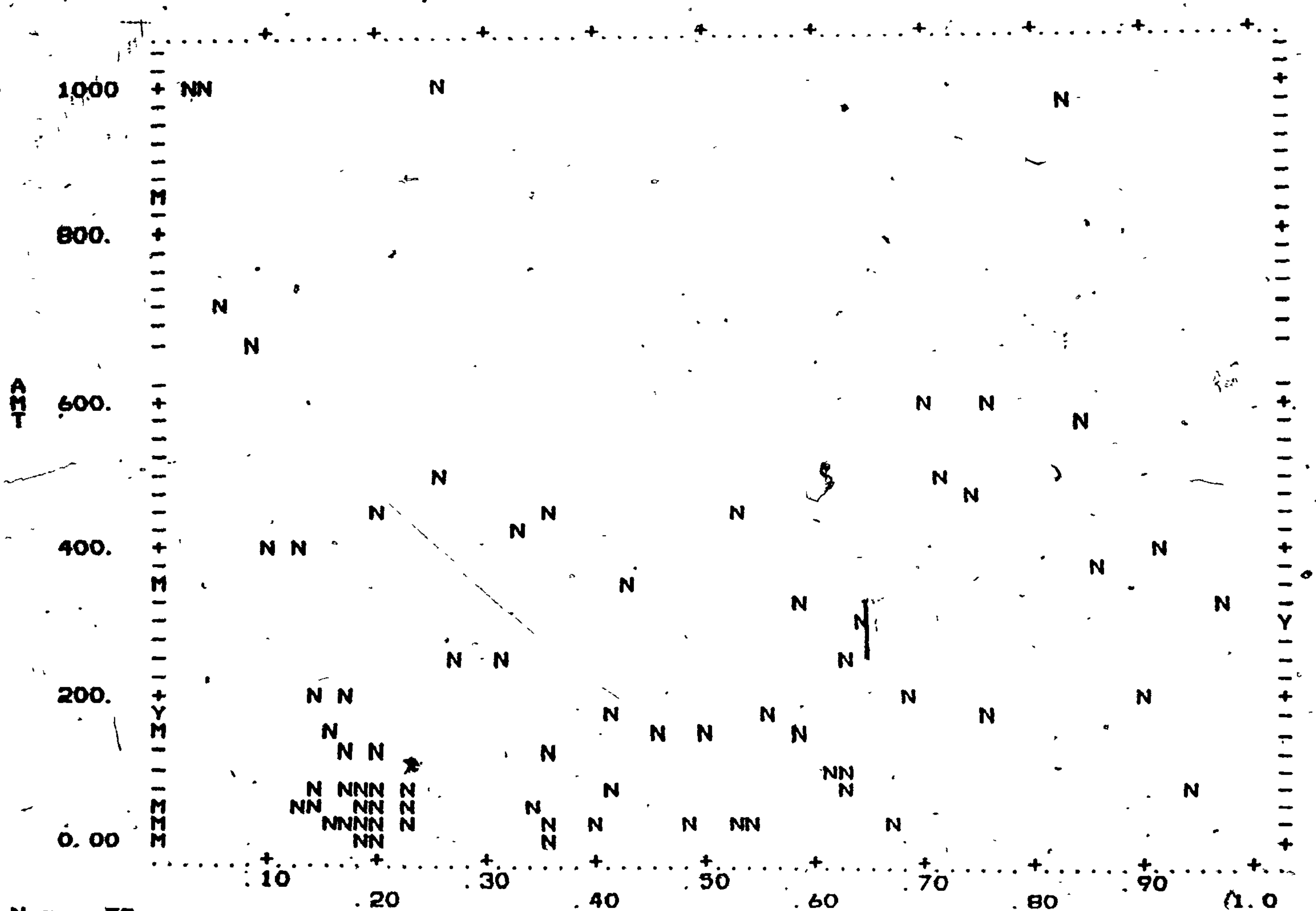
LOGP VERSUS AMTTN (28 VS. 31) GROUP=YES , SYMBOL=Y
1PAGE 26 BMDP6D PLOTS ANAL OF BDATA




```
N = 283
R = .0376
P = .5294

--REGRESSION LINE--      --RMS--         MEAN      S.D.
Y= 1.0808 +3.0264*X       891.91     X   .72097    .37032
                                     Y  3.2627    29.833

LOOP     VERSUS AMTTN    ( 28 VS. 31 )   GROUP=NO     , SYMBOL=N
                                         GROUP=YES    , SYMBOL=Y
1PAGE  27  BMDP6D PLOTS ANAL OF BDATA
```

TABLE A-6.1.2

LRME

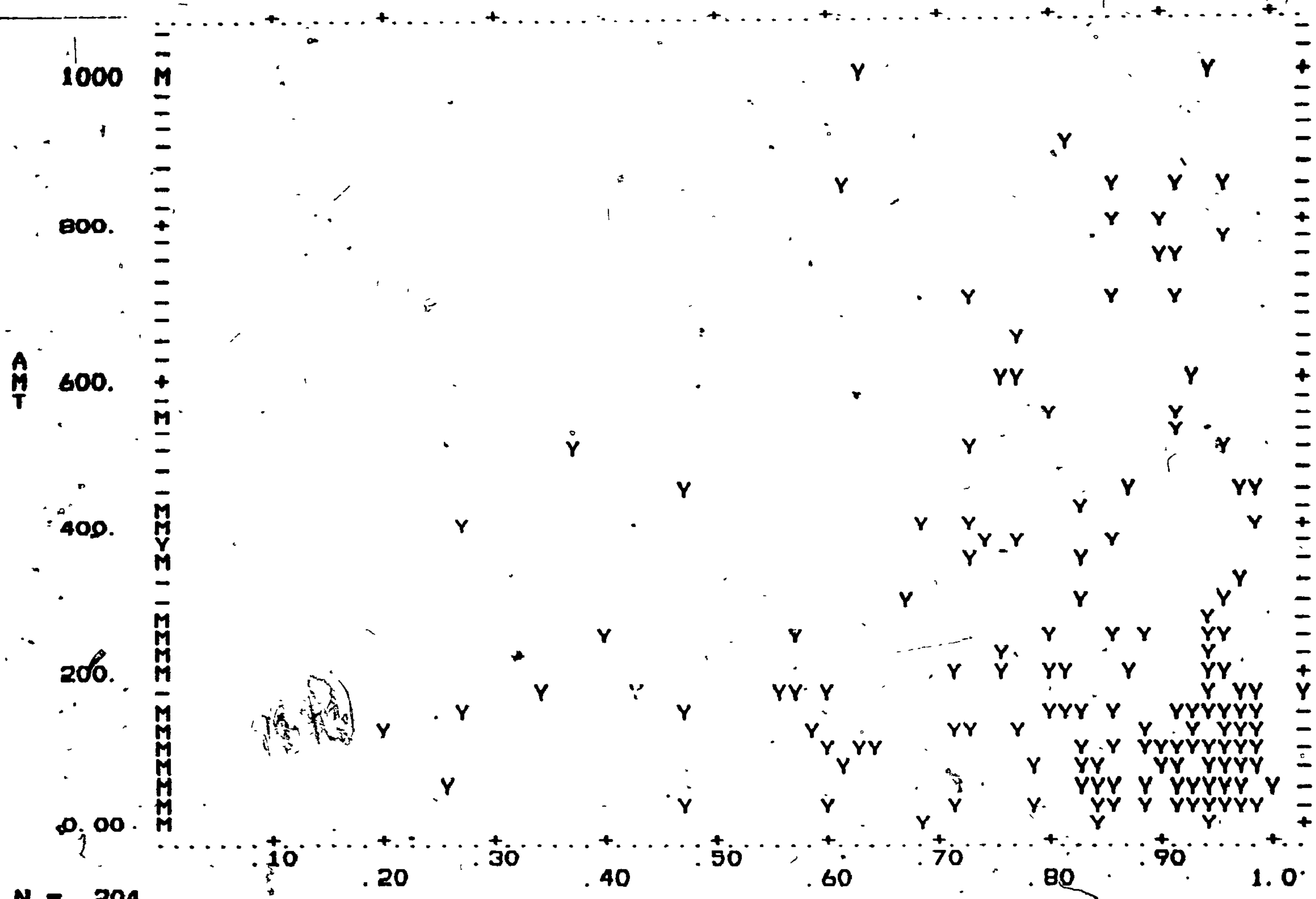


LRPP

N = 79
R = .1254
P = .2717

--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y= 180.50 +124.55*X	64408.	X	.38715	.25591
		Y	228.72	254.16

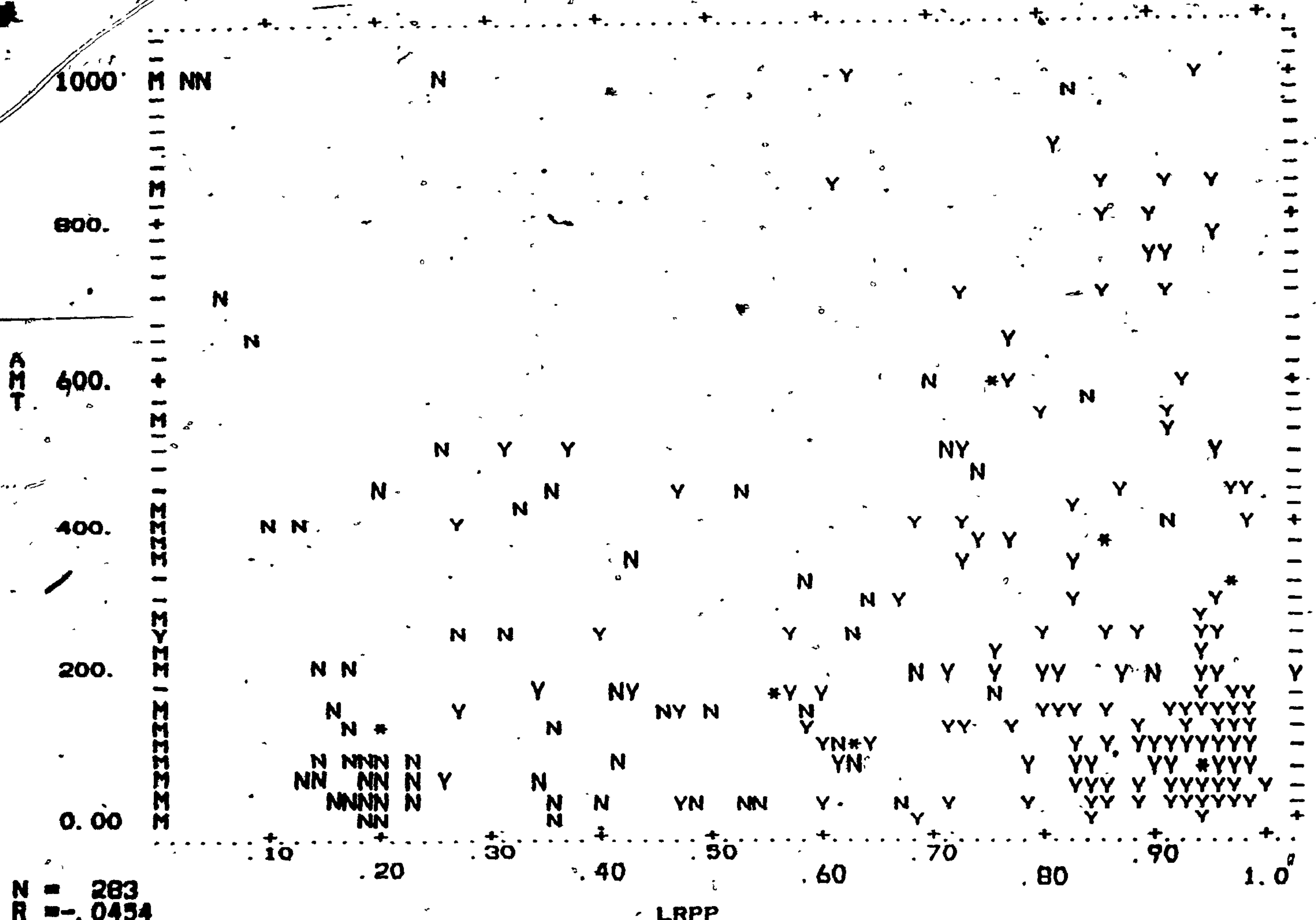
LRPP VERSUS AMT (29 VS. 3) . GROUP=NO , SYMBOL=N
1PAGE 7 BMDP6D PLOTS ANAL OF BDATA



N = 204
R =-.1471
P = .0357

--REGRESSION LINE--	--RMS--		MEAN	S.D.
		X	.84994	.16336
Y= 385.35 -201.62*X	49297.	Y	213.99	223.92

LRPP VERSUS AMT (29 VS. 3) GROUP=YES , SYMBOL=Y
1PAGE 8 BMDP6D PLOTS ANAL OF BDATA

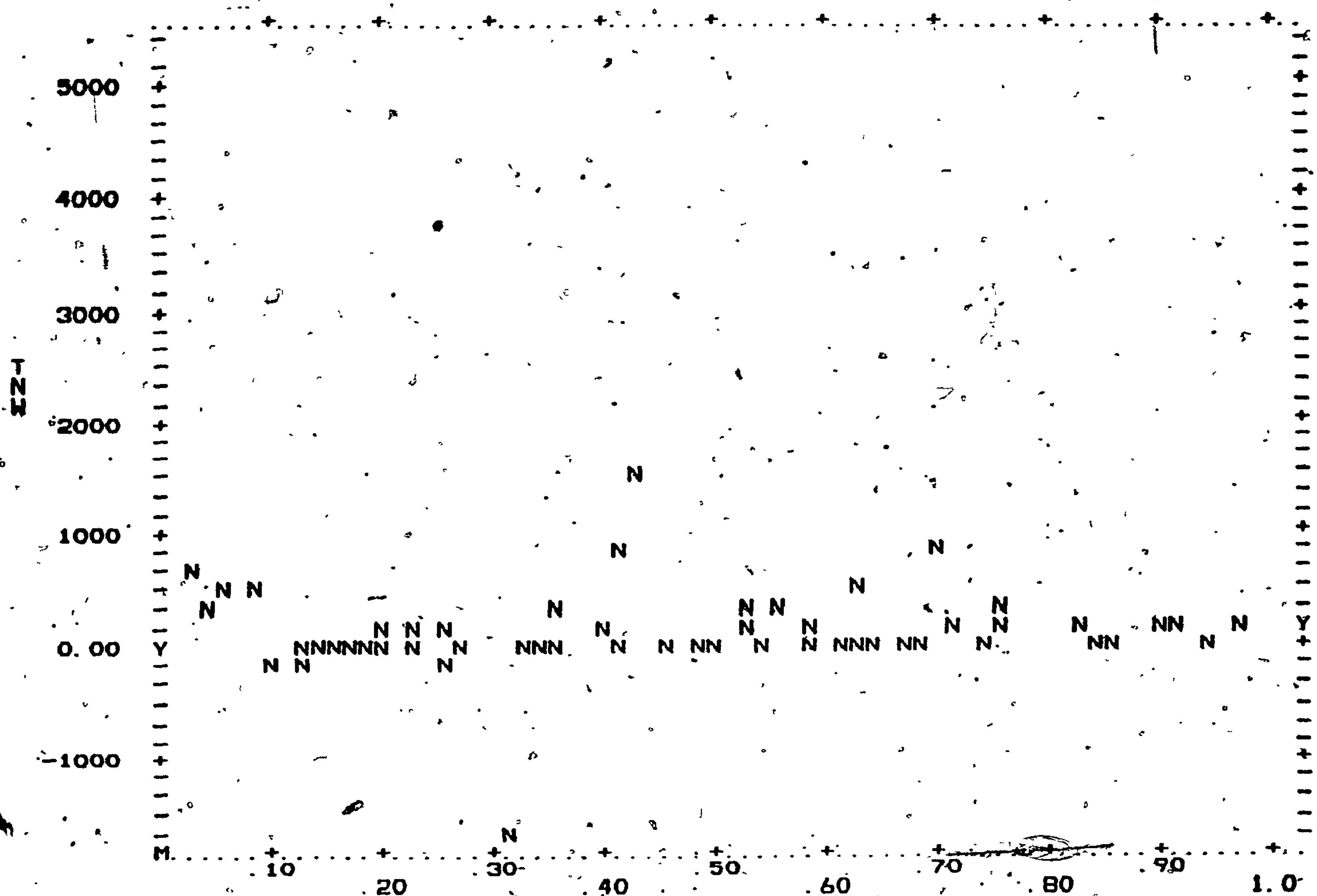


N = 283
R =-.0454
P = .4473

--REGRESSION LINE--	--RMS--		MEAN	S. D.
Y= 244.88 -37.155*X	54085.	X	.72075	.28386
		Y	218.10	232.39

LRPP VERSUS AMT (29 VS. 3) GROUP=NO ; SYMBOL=N
GROUP=YES ; SYMBOL=Y

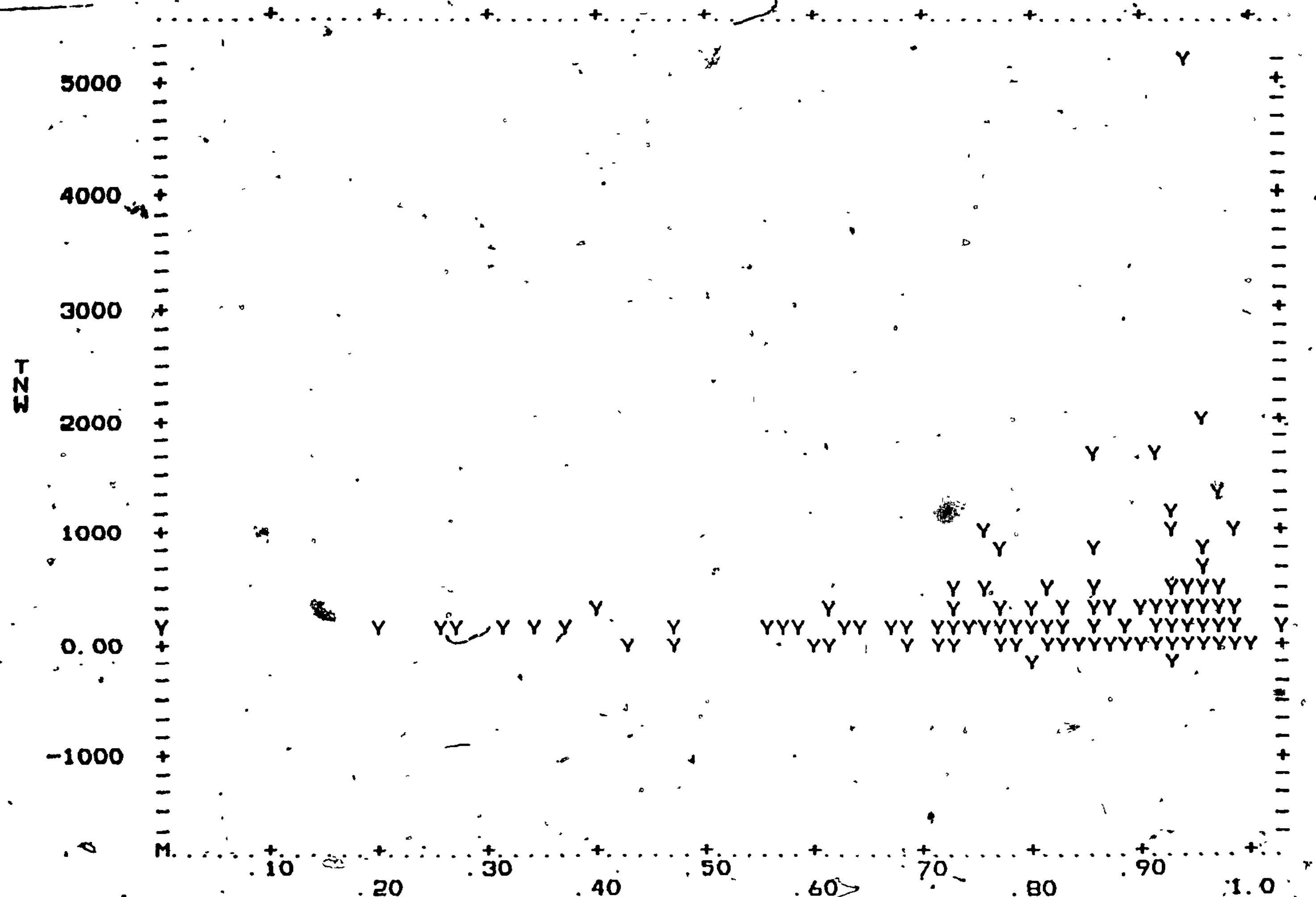
1PAGE 9 BMDP6D PLOTS ANAL OF BDATA



N = .79
R = .1138
P = .3188

--REGRESSION LINE--	--RMS--		MEAN	S. D.
Y= 42.958 +139.39*X	98161.	X	.38715	.25591
		Y	96.924	313.33

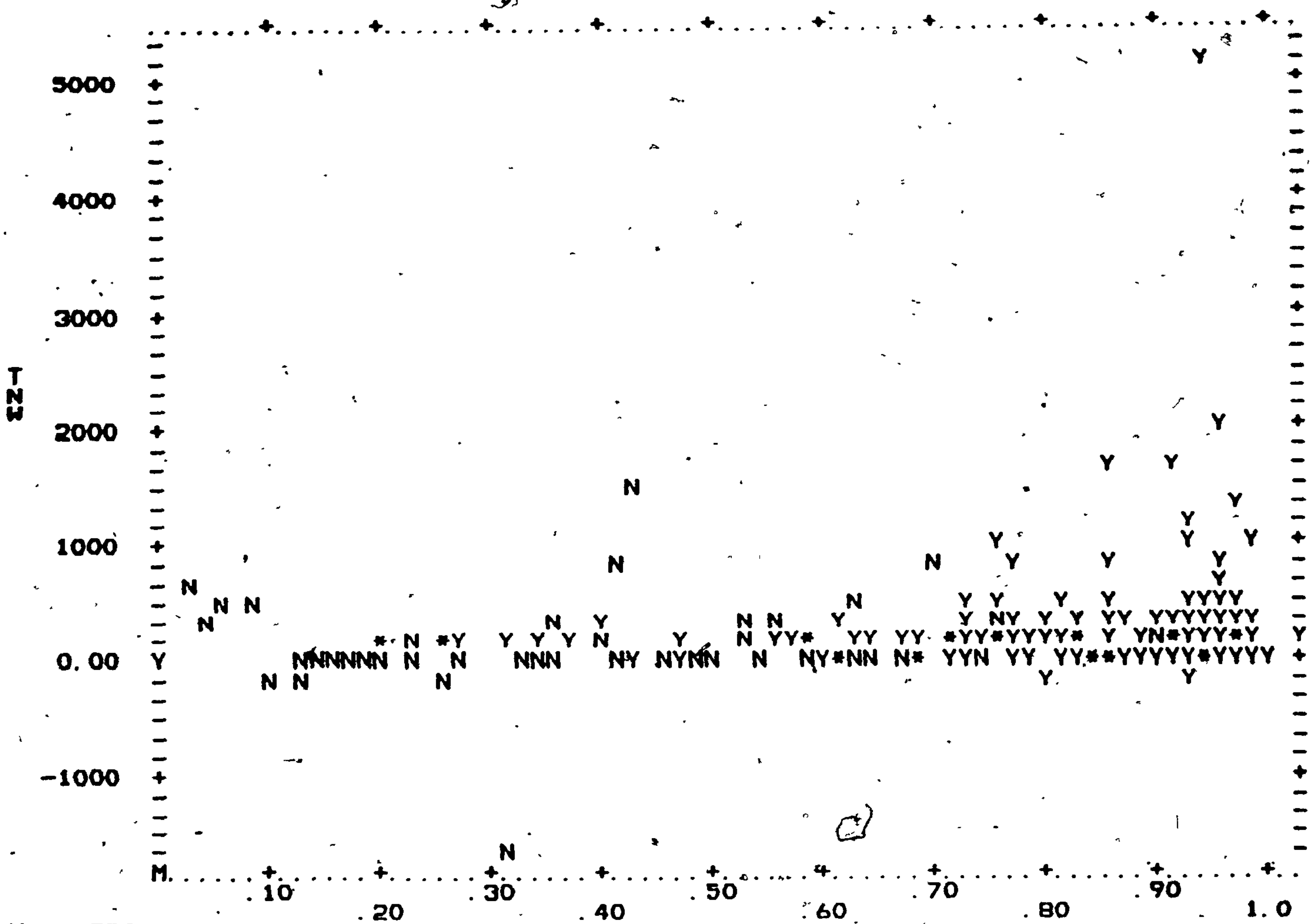
LRPP VERSUS TNW (29 VS. 19) GROUP=NO , SYMBOL=N
1PAGE 19 BMDP6D PLOTS ANAL OF BDATA



N = 204
R = .0474
P = .5016

			MEAN	S. D.
--REGRESSION LINE--	--RMS--	X	.84994	.16336
Y= 93.481 +131.40*X	205922	Y	205.16	453.18

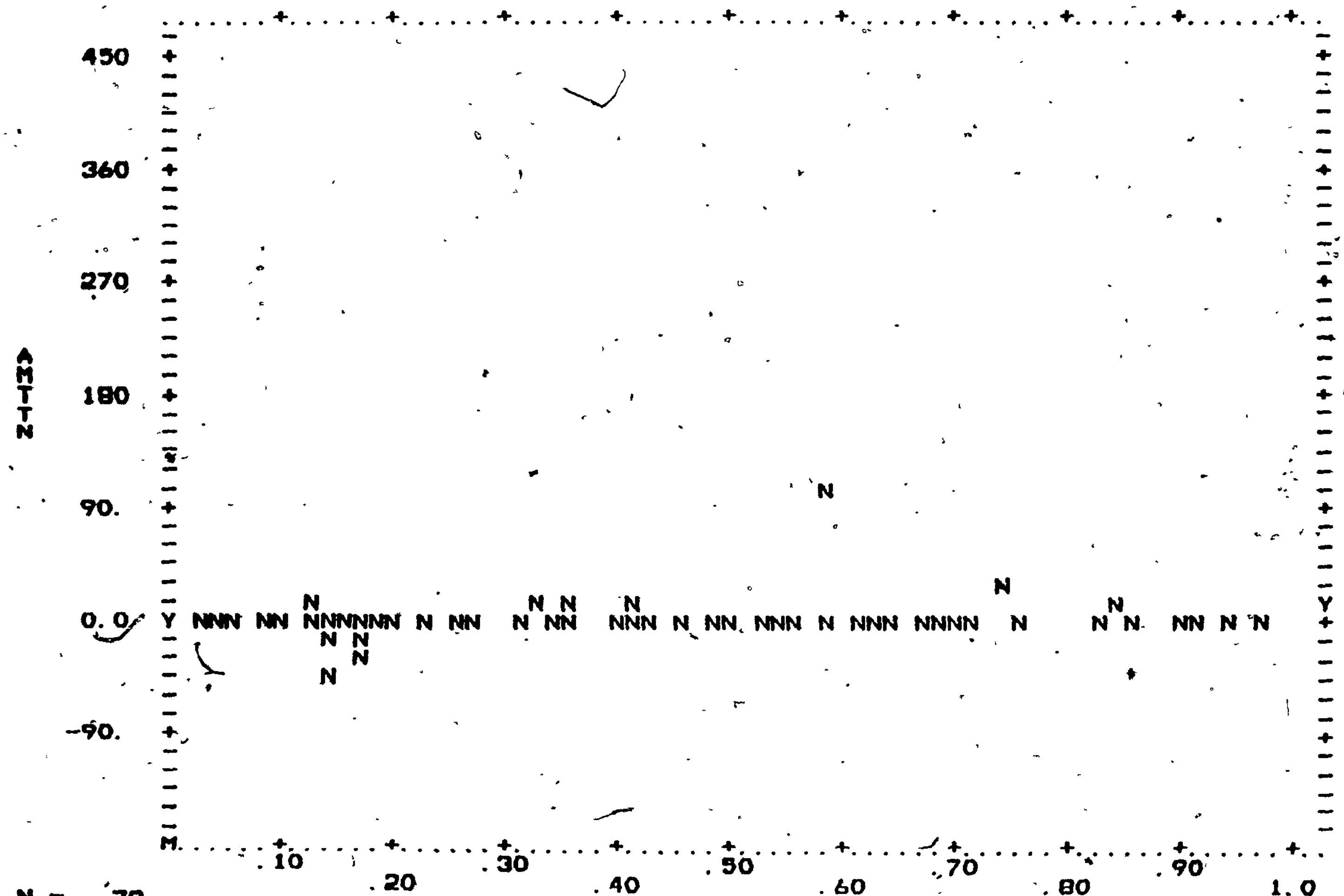
LRPP VERSUS TNW (29 VS. 19) GROUP=YES SYMBOL=Y
1PAGE. 20 BMDP6D PLOTS ANAL OF BDATA



N = 283
R = .1269
P = .0328

--REGRESSION LINE--	--RMS--		MEAN	S. D.
Y= 39.298 +188.20*X	175123	X	.72075	.28386
		Y	174.95	421.14

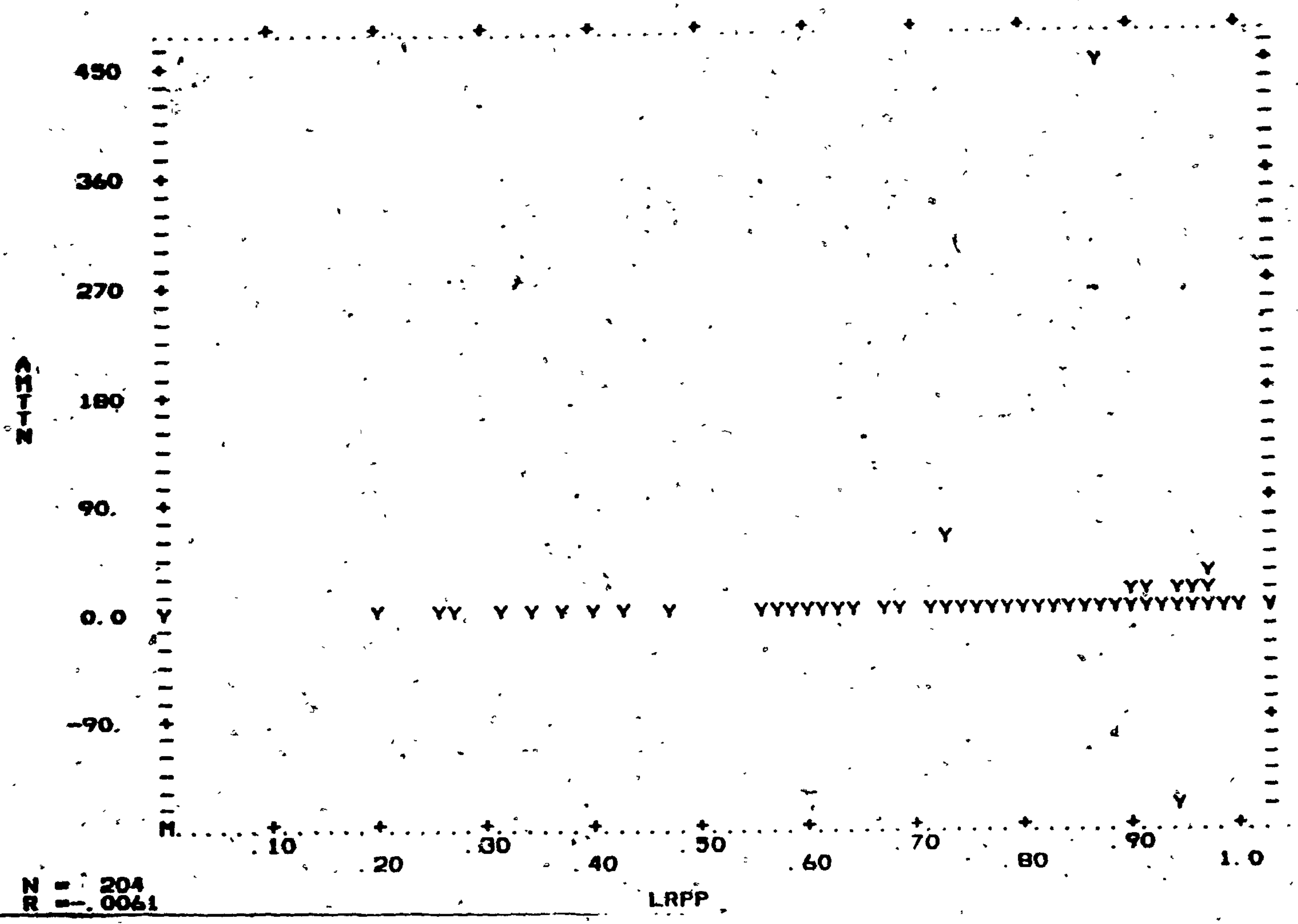
LRPP VERSUS TNW (29 VS. 19) GROUP=NO , SYMBOL=N
GROUP=YES , SYMBOL=Y



N = 79
R = .2365
P = .0356

--REGRESSION LINE--	--RMS--		MEAN	S.D.
		X	.38715	.25591
Y=-3.4971 +13.638*X	208.32	Y	1.7827	14.759

LRPP VERSUS AMTTN (29 VS. 31) GROUP=NO SYMBOL=N
1PAGE 25 BMDP6D PLOTS ANAL OF BDATA

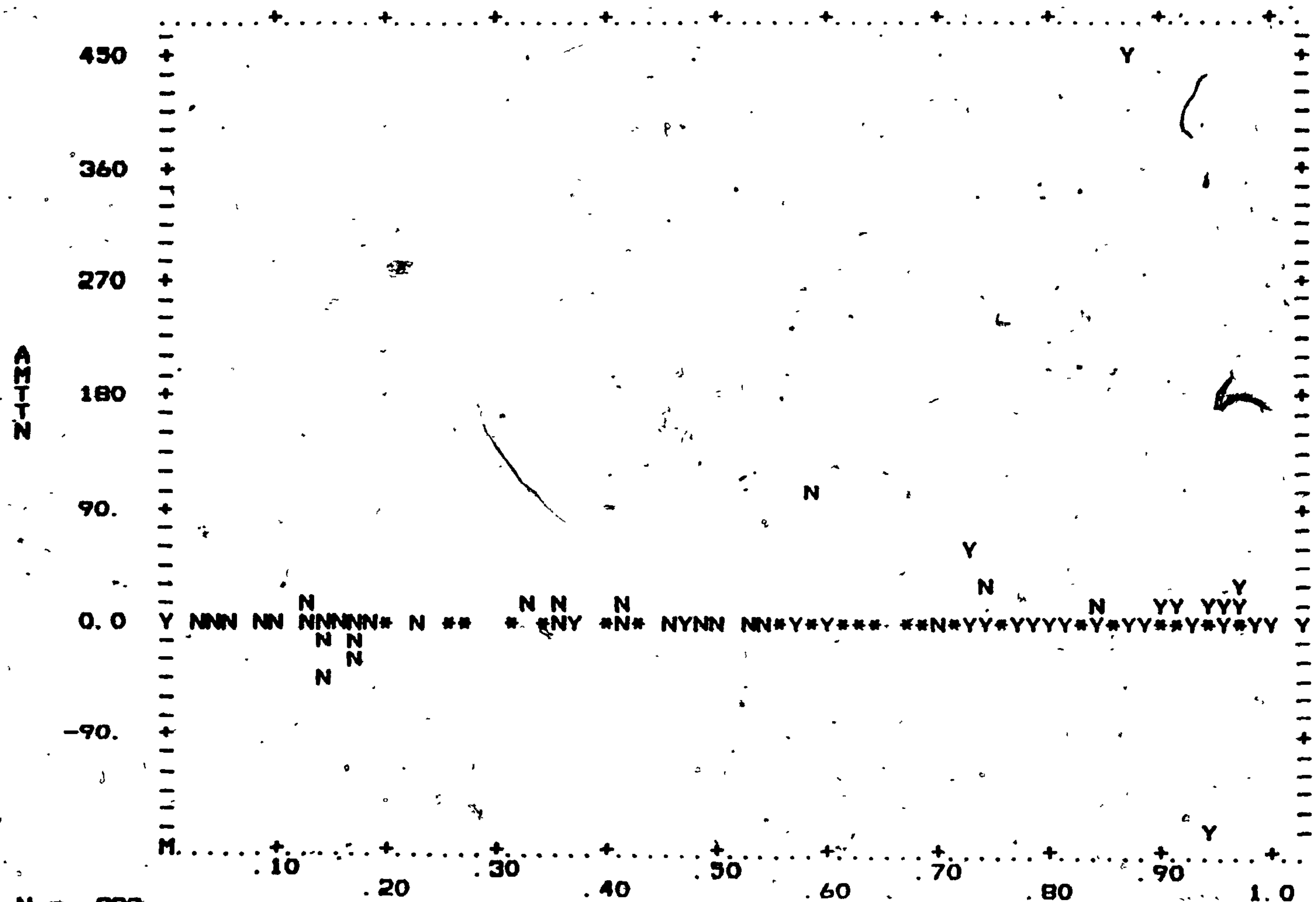


N = 204
R =-.0061

P = .9305

--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y= 4.9216 -1.2774*X	1157.1	X	.84994	.16336
		Y	3.8359	33.933

LRPP VERSUS AMTTN (29 VS. 31) GROUP=YES , SYMBOL=Y
1PAGE 26 BMDP6D PLOTS ANAL OF BDATA



N = 283

R = .0489
P = .4125

LRPP

--REGRESSION LINE--	--RMS--		MEAN	S.D.
Y=-.44394 +5.1428*X	891.03	X	.72075	.28386
		Y	3.2627	29.833

LRPP VERSUS AMTTN (29 VS. 31) GROUP=NO ; SYMBOL=N
GROUP=YES ; SYMBOL=Y

1PAGE 27 BMDP6D PLOTS ANAL OF BDATA

TABLE A-6.2

TABLE A-6.2.1

LRFE

PAGE 4 BMDP7D PLOTS ANAL OF BDATA

HISTOGRAM OF * AMT * (3) GROUPED BY * LOOP * (28)
AND * OUT * (2)

MIDPOINTS	.200000 NO	.200000 YES	.400000 NO	.400000 YES	.600000 NO	.600000 YES	.800000 NO	.800000 YES	LAST NO	LAST YES
1200.000)										
1150.000)										
1100.000)										
1050.000)										
1000.000)	***	*		*					*	*
950.000)										
900.000)										*
850.000)	*					*				***
800.000)								*		**
750.000)										**
700.000)	*									***
650.000)	*									*
600.000)							***			***
550.000)		*							N	**
500.000)	*	*		*	*		*	*		****
450.000)		*	**				*	*		*****
400.000)	****	**					N		*	****
350.000)	**	**	*							*****
300.000)		*		N	*	M	*			****
250.000)	*	*	*				*	N		*********11
200.000)	M**	**	N	**	M	*	*	****	*	M********11
150.000)	****	M**	**	**	*		*	**		*********22
100.000)	*********	*******			***	*		*		*********42
50.000)	*********18	*********18	***	**		*		*		*********39
.000)	*********11	*********11	*			*		*		*********12

GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE

	.200000 NO	.200000 YES	.400000 NO	.400000 YES	.600000 NO	.600000 YES	.800000 NO	.800000 YES	LAST NO	LAST YES
MEAN	189.966	136.843	195.100	282.500	198.857	289.600	398.444	267.000	528.333	203.328
STD. DEV.	265.872	186.590	171.003	322.988	155.806	329.497	181.480	224.243	403.859	216.747
S. E. M	34.911	26.128	54.076	114.194	58.889	147.356	60.493	67.612	233.168	16.292
MAXIMUM	1000.000	1000.000	450.000	1000.000	500.000	850.000	602.000	800.000	980.000	1000.000
MINIMUM	7.000	5.000	20.000	35.000	78.000	35.000	150.000	25.000	202.000	8.000
CASES INCL	58	51	10	8	7	5	9	11	3	177

ALL GROUPS COMBINED
(EXCEPT CASES WITH UNUSED VALUES
FOR VARIABLES LOOP AND OUT)

MEAN	203.968
STD. DEV	229.046
S. E. M	12.440
MAXIMUM	10.0E+2
MINIMUM	5.000
CASES EXCLUDED	(0)
CASES INCLUDED	339

ROBUST S.D 215.047

```
PAGE   6  BMDP7D PLOTS ANAL OF BDATA

                    ************                         ************
HISTOGRAM OF  * TNW        *  (  19)   GROUPED  BY   * LOOP       *  (   28)
                    ************                AND  * OUT        *  (    2)
                                                     ************

               * 200000   * 200000   * 400000   * 400000   * 600000   * 600000   * 800000   * 800000   *LAST      *LAST
               NO         YES        NO         YES        NO         YES        NO         YES        NO         YES
MIDPOINTS      + . . . .  + . . . .  + . . . .  + . . . .  + . . . .  + . . . .  + . . . .  + . . . .  + . . . .  + . . . . +
5400. 000)
5100. 000)                                                                                                        *
4800. 000)
4500. 000)
4200. 000)
3900. 000)
3600. 000)
3300. 000)
3000. 000)
2700. 000)
2400. 000)
2100. 000)                                                                                                        *
1800. 000)
1500. 000)                                                                                                        ***
1200. 000)                                                                                                        ***
 900. 000)*                                                                      *                                ****
 600. 000)***                                                                    *          *                     *****
 300. 000)**          *          M*         ***        **         *          M***       M****      M**        M********38
    . 000)M********43.M*                    *******    M****      M****      M***       ***        *****                 ********121
-300. 000)                                                                                                        *
-600. 000)
-900. 000)
*********)
*********)
*********)
          GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE

MEAN            37 840    132 333    219. 300    110 875     90 143    124. 200    263. 889    209 455    188. 667    212. 678
STD. DEV.      305. 085    63. 956    460. 620     59. 266    112. 863    137. 520    252. 448    149. 799    12. 220    483. 987
S. E. M         43 145     36. 925    145 661     20 954     42. 658     61 501     84. 149     45 166     7. 055     36. 379
MAXIMUM        807. 000    206. 000   1499. 000    180 000    313. 000    361. 000    813 000    483 000    202. 000   5100. 000
MINIMUM      -1659. 000     91. 000    -36. 000     26. 000    -25. 000     16. 000     17 000     38. 000    178. 000   -179. 000
CASES EXCL     (    8)    (   48)    (    0)    (    0)    (    0)    (    0)    (    0)    (    0)    (    0)    (    0)
CASES INCL         50          3         10          8          7          5          9         11          3        177.

                 ALL GROUPS COMBINED
              (EXCEPT CASES WITH UNUSED VALUES
            FOR VARIABLES LOGP        AND OUT         )
                      MEAN              174 947
                      STD  DEV          421 137
                      S. E. M            25 034
                      MAXIMUM            5. 1E+3
                      MINIMUM           -1 7E+3
                      CASES EXCLUDED    (   56)
                      CASES INCLUDED       283

                      ROBUST S D        250 451
```

```
PAGE   6  BMDP7D PLOTS ANAL OF BDATA

            ************                        ************
HISTOGRAM OF * AMTTN      *  (  31)   GROUPED  BY   * LOCP      *  (  28)
            ************                  AND   * OUT       *  (   2)
                                                ************

            * 200000   * 200000   * 400000   * 400000   * 600000   * 600000   * 800000   * 800000   *LAST     *LAST
            NO         YES        NO         YES        NO         YES        NO         YES        NO        YES
MIDPOINTS
  510.000)
  480.000)
  450.000)
  420.000)
  390.000)
  360.000)
  330.000)
  300.000)
  270.000)
  240.000)
  210.000)
  180.000)
  150.000)
  120.000)
   90.000)
   60.000)                                                                                                   *
   30.000)*                                          N*                    *                                 **
     .000)M*******46 M**        M********* M*******  ******     M****      M*******   M*******11 M**          M******172
  -30.000)**
  -60.000)*
  -90.000)
 -120.000)
 -150.000)
 -180.000)
 -210.000)
         GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE
```

MEAN	-0.960	2.753	2.110	2.635	15.352	2.238	5.787	1.566	2.734	4.095
STD DEV	10.014	1.970	4.615	2.197	39.631	0.398	8.953	1.307	1.936	36.431
S E M	1.416	1.138	1.459	0.777	14.979	0.178	2.984	0.394	1.112	2.738
MAXIMUM	19.091	5.000	10.000	6.135	105.000	2.857	28.235	4.605	4.851	450.000
MINIMUM	-50.000	1.319	-6.972	0.427	-5.600	1.887	0.545	0.269	1.086	-166.000
CASES EXCL	(8)	(48)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
CASES INCL	50	3	10	8	7	5	9	11	3	177

```
ALL GROUPS COMBINED
(EXCEPT CASES WITH UNUSED VALUES
FOR VARIABLES LOCP       AND OUT        )
     MEAN                  3.263
     STD  DEV             29.833
     S  E  M               1.773
     MAXIMUM             450.000
     MINIMUM              -1.7E+2
     CASES EXCLUDED       (   56)
     CASES INCLUDED          283

     ROBUST S D            7.117
```

TABLE A-6.2.2

LRME

```
PAGE    -  BMDP7! PLOTS ANAL OF BDATA

                    ************                                ************
HISTOGRAM OF      * AMT        *  (    3)    GROUPED  BY        * LRPP       *  (   29)
                    ************                      AND       * OUT        *  (    2)
                                                                ************

                                                                                                                                          CASES
                                                                                                                                            UNUS
                                                                                                                                          VALUES
          * .200000     * .200000     * .400000     * .400000     * .600000     * .600000     * .800000     * .800000     *LAST         *LAST         LRPP
            NO            YES           NO            YES           NO            YES           NO            YES           NO            YES           OUT
MIDPOINTS    +             +             +             +             +             +             +             +             +             +             +
 1200.000)
 1150.000)
 1100.000)
 1050.000)
 1000.000)**                          *                                                                           *             *             *             *
  950.000)
  900.000)                                                                                                                                    *
  850.000)                                                                                                        *                           ***           *
  800.000)                                                                                                                                    ***
  750.000)                                                                                                                                    **
  700.000)                                                                                                        *                           **
  650.000)                                                                                                        *
  600.000)                                                                                          **            **            *             *
  550.000)                                                                                                        *                           *             *
  500.000)                            *             **                                              **            **                          ***
  450.000)*                           *                           *             *                                                             *****         *
  400.000)**                          *             *                                                             **            M*            **            *
  350.000)                                                        *                                               ***           *             **            ***
  300.000)                                          N             *                             *             M                             ****          *
  250.000)                            **                                        **              M             **                            *******       *
  200.000)M*                          N             *             *             M**             **            ****          *             M**********   **
  150.000)***                                       *             M***          **                            **                          *********21   M***
  100.000)****          M             **                          *                             ****          ********                    ********36    *****
   50.000)*********11                 *****         *             *             *               *             ***           *             *********38   *****
      000)*****                       **                          **                                          **                          **********    *****
         GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE

MEAN         196 156      120 000      221.667      297 500      169.500      201 667      268 923      295 361      419 571      192 770      132
STD. DEV     277 779        0 000      273 834      194 004      137 370      113 285      207 737      253.085      294 020      219.672      198
S  E  M       49 105        0 000       70 704       79 202       39 655       37 762       57 616       42.181      111.129       17 818       26
MAXIMUM     1000.000      120 000     1000.000      500 000      450 000      450 000      602 000     1000 000      980 000     1000 000     1000
MINIMUM       10 000      120 000       11 000       50 000       15 000       35 000       25 000        8.000       70 000       10 000        5
CASES INCL        32            1           15            6           12            9           13           36            7          152

                    ALL GROUPS COMBINED
              (EXCEPT CASES WITH UNUSED VALUES
          FOR VARIABLES LRPP       AND OUT          )
                    MEAN                 218 162
                    STD.  DEV.           232 388
                    S  E  M               13 814
                    MAXIMUM             10.0E+2
                    MINIMUM                5.000
                    CASES EXCLUDED      (      0)
                    CASES INCLUDED           283

                    ROBUST S-D           266 446
```

PAGE 6 BMDP7D PLOTS ANAL OF BDATA

```
                   ************                       ************
HISTOGRAM OF       * TNW      *  ( 19)   GROUPED BY   * LRPP     *  ( 29)
                   ************                 AND   * OUT      *  (  2)
                                                      ************

             * 200000  * 200000  * 400000  * 400000  * 600000  * 600000  * 800000  * 800000  *LAST     *LAST      CASES
              NO        YES       NO        YES       NO        YES       NO        YES       NO        YES       UNUS
                                                                                                                  VALUES
                                                                                                                  LRPP
                                                                                                                  OUT
MIDPOINTS    +.........+.........+.........+.........+.........+.........+.........+.........+.........+.........+.....
 5400.000)
 5100.000)                                                                                              *
 4800.000)
 4500.000)
 4200.000)
 3900.000)
 3600.000)
 3300.000)
 3000.000)
 2700.000)
 2400.000)
 2100.000)                                                                                              *
 1800.000)
 1500.000)                                       *                                                      ***
 1200.000)                                                                                              ***
  900.000)                                       *                   *         **                       **
  600.000)***                                                        *         **                       ****
  300.000)**                  **       ***       M**       M***      M**       M*********     ***       M********32
     .000)M*******27 M        M*******12 M**     *******   *****     *********  ********23 M***         ******105
 -300.000)                                                                                              *
 -600.000)
 -900.000)
********)
********)
********)
         GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE
```

MEAN	67.375	91.000	-59.000	146.667	276.667	151.111	174.615	184.056	113.714	216.421	0
STD. DEV.	180.301	0.000	453.199	45.798	449.061	116.961	233.559	229.689	82.504	512.287	0
S. E. M	31.873	0.000	117.016	18.697	129.633	38.987	64.778	38.281	31.184	41.552	0
MAXIMUM	709.000	91.000	300.000	206.000	1499.000	405.000	813.000	998.000	202.000	5100.000	0
MINIMUM	-137.000	91.000	-1659.000	100.000	-25.000	26.000	17.000	-148.000	-28.000	-179.000	0
CASES INCL	32	1	15	6	12	9	13	36	7	152	

```
        ALL GROUPS COMBINED
  (EXCEPT CASES WITH UNUSED VALUES
FOR VARIABLES LRPP     AND OUT       )
       MEAN            174.947
       STD  DEV        421.137
       S  E  M          25.034
       MAXIMUM          5.1E+3
       MINIMUM         -1.7E+3
       CASES EXCLUDED   (   0)
       CASES INCLUDED      283

       ROBUST S D      250.451
```

PAGE 6 BMDP7I FLOTS ANAL OF BDATA

```
            ************                                ************
HISTOGRAM OF * AMTTN    * ( 31)   GROUPED  BY      * LRPP      * ( 29)
            ************                AND      * OUT       * ( 2)
                                                 ************

          * 200000  * 200000  * 400000  * 400000  * 600000  * 600000  * 800000  * 800000  *LAST  *LAST   CASES UNUSED
            NO        YES       NO        YES       NO        YES       NO        YES      NO     YES    VALUES LRPP OUT
MIDPOINTS
 510.000)
 480.000)
 450.000)
 420.000)
 390.000)
 360.000)
 330.000)
 300.000)
 270.000)
 240.000)
 210.000)
 180.000)
 150.000)
 120.000)
  90.000)
  60.000)
  30.000)
    .000)M********29 M    M********14 M****   M********11 M*********  M********12 M********35 M******   M*******148
 -30.000)**
 -60.000)*
 -90.000)
-120.000)
-150.000)
-180.000)
-210.000)
GROUP MEANS ARE DENOTED BY M'S IF THEY COINCIDE WITH *'S, N'S OTHERWISE
```

	* 200000 NO	* 200000 YES	* 400000 NO	* 400000 YES	* 600000 NO	* 600000 YES	* 800000 NO	* 800000 YES	*LAST NO	*LAST YES	CASES UNUSED VALUES LRPP OUT
MEAN	-2 848	1 319	2 457	2 360	9 691	1 800	4 361	3 432	3 160	4 127	0
STD DEV	11 651	0 000	5 944	2 067	30 168	1 240	7 420	10 445	3 487	39 013	0
S E M	2 060	0 000	1 535	0 844	8 709	0 413	2 058	1 741	1 318	3 164	0
MAXIMUM	8 524	1 319	19 091	5 000	105 000	4 512	28 235	63 636	8 612	450 000	0
MINIMUM	-50 000	1 319	-6 972	0 427	-5 600	0 617	0 545	-3 716	-2 500	-166 000	0
CASES IN	32	1	15	6	12	9	12	35	7	152	

ALL GROUPS COMBINED
(EXCEPT CASES WITH UNUSED VALUES
FOR VARIABLES LRPP AND OUT)

MEAN	3 263
STD DEV	29 833
S E M	1 773
MAXIMUM	450 000
MINIMUM	-1 7E+2
CASES EXCLUDED	(0)
CASES INCLUDED	283
ROBUST S D	3 117

TABLE A-6.3

TABLE A-6.3.1

LDA

PAGE 11 BMDP4F TABLES ANAL OF BDATA

```
************************
* TABLE PARAGRAPH   1  *
************************
```

OBSERVED FREQUENCY TABLE 1

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	37	0	1	0	38
	.400000	4	1	0	7	12
	.600000	8	1	8	1	18
	.800000	0	2	0	0	2
	LAST	4	6	0	7	17
	TOTAL	53	10	9	15	87
YES	.200000	8	0	0	1	9
	.400000	0	1	0	5	6
	.600000	9	7	3	4	23
	.800000	0	2	0	1	3
	LAST	68	34	39	70	211
	TOTAL	85	44	42	81	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

ALL CASES HAD COMPLETE DATA FOR THIS TABLE.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	39	816.53	.00000	1327.43	.00000
1	31	271.16	.00000	382.44	.00000
2	8	26.14	.00099	21.50	.00593
3	0	0.	1	0	1

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	8	545.37	.00000	944.99	.00000
2	23	245.02	.00000	360.94	.00000
3	8	26.14	.00099	21.50	.00593

PAGE 13 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 2

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	0	38	0	0	38
	.400000	3	7	0	2	12
	.600000	14	1	1	2	18
	.800000	0	1	0	1	2
	LAST	2	13	1	1	17
	TOTAL	19	60	2	6	87
YES	.200000	2	5	2	0	9
	.400000	0	6	0	0	6
	.600000	9	5	4	5	23
	.800000	1	2	0	0	3
	LAST	70	81	19	41	211
	TOTAL	82	99	25	46	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

ALL CASES HAD COMPLETE DATA FOR THIS TABLE.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D. F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	39	862.06	.00000	1502.77	.00000
1	31	251.75	.00000	336.76	.00000
2	9	31.24	.00027	28.78	.00071
3	0	0.	1	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D. F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	610.30	.00000	1166.01	.00000
2	22	220.52	.00000	307.98	.00000
3	9	31.24	.00027	28.78	.00071

		1 00000	2 00000	4 00000	5 00000	6 00000	10 0000	11 0000	12 0000	14 0000	15 0000	17.0000	18 0000
NO	200000	0	0	7	0	0	0	27	0	0	0	0	0
	400000	0	0	1	0	0	0	6	0	0	0	0	0
	.600000	0	0	9	0	0	0	8	0	0	0	0	0
	800000	0	0	0	0	0	0	0	0	0	0	0	0
	LAST	3	2	0	0	1	0	0	0	0	0	0	0
	TOTAL	3	2	17	0	1	0	41	0	0	0	0	0
YES	.200000	0	0	1	0	0	0	4	0	0	0	0	0
	400000	0	0	0	0	0	0	5	0	0	0	0	0
	600000	0	0	[illegible]	0	0	0	11	0	0	0	0	0
	800000	0	0	0	0	0	0	1	0	0	0	0	0
	LAST	22	11	0	2	9	4	3	3	2	6	4	[illegible]
	TOTAL	22	11	[illegible]	2	9	4	24	3	2	6	4	[illegible]

OUT LDAP BRN

		19 0000	21 0000	22 0000	24 0000	26 0000	27 0000	28 0000	30 0000	31 0000	32 0000	33 0000	34 0000
NO	200000	4	0	0	0	0	0	0	0	0	0	0	0
	.400000	5	0	0	0	0	0	0	0	0	0	0	0
	.600000	0	0	0	0	0	0	0	0	0	0	0	0
	800000	0	0	0	0	0	1	0	0	0	0	0	0
	LAST	0	1	0	0	0	1	0	1	0	0	1	1
	TOTAL	9	1	0	0	0	2	0	1	0	0	1	1
YES	200000	4	0	0	0	0	0	0	0	0	0	0	0
	400000	0	0	0	0	0	0	0	0	0	0	0	0
	600000	0	0	0	0	0	0	0	0	0	0	0	1
	800000	0	0	0	0	0	0	0	0	0	0	0	0
	LAST	0	14	1	11	4	16	1	5	11	2	3	0
	TOTAL	4	14	1	11	4	16	1	5	11	2	3	1

OUT LDAP BRN

		35 0000	36 0000	37 0000	38 0000	39 0000	40 0000	41 0000	42 0000	43 0000	44.0000	45 0000	46 0000
NO	200000	0	0	0	0	0	0	0	0	0	0	0	0
	400000	0	0	0	0	0	0	0	0	0	0	0	0
	600000	0	0	0	1	0	0	0	0	0	0	0	0
	800000	0	0	0	0	0	0	0	0	0	0	0	0
	LAST	0	1	1	0	0	1	0	1	0	0	0	1
	TOTAL	0	1	1	1	0	1	0	1	0	0	0	1
YES	200000	0	0	0	0	0	0	0	0	0	0	0	0
	400000	0	0	0	0	0	0	0	0	0	0	0	0
	600000	0	0	0	0	0	0	0	0	0	0	0	1
	800000	0	0	0	0	0	0	0	1	0	0	0	
	LAST	4	7	5	0	1	16	5	3	6	3	2	
	TOTAL	4	7	5	0	1	16	5	4	6	3	2	

OUT LDAP BRN

		47 0000	48 0000	TOTAL
NO	200000	0	0	38
	.400000	0	0	12
	.600000	0	0	18
	.800000	1	0	2
	LAST	1	0	17
	TOTAL	2	0	87
YES	.200000	0	0	9
	.400000	1	0	6
	.600000	0	0	23
	.800000	1	0	3
	LAST	0	3	201
	TOTAL	2	3	242

TOTAL OF THE OBSERVED-FREQUENCY TABLE IS 325

10 CASES HAD INCOMPLETE DATA.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	379	1410.02	.00000	3548.39	.00000
1	337	611.97	.00000	974.42	.00000

** THE LOGLINEAR ALGORITHM DID NOT CONVERGE WHEN FITTING THE MODEL BELOW.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
2	9	26.07	.001[illegible]	22.37	.00778
3	0	0.	1	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	42	798.05	.00000	2573.96	.00000
2	328	585.90	.00000	952.05	.00000
3	9	26.07	.001[illegible]	22.37	.00778

PAGE 18 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 4

OUT	LDAP	TYNEW 0 00000	1.00000	2.00000	3 00000	4 00000	5.00000	6 00000	7 00000	8.00000	9 00000	10 0000	11 0000
NO	.200000	0	0	0	0	0	37	0	1	0	0	0	0
	.400000	0	2	1	0	0	0	0	0	7	0	0	0
	.600000	0	5	1	7	1	1	0	0	0	1	0	0
	.800000	0	0	0	0	0	0	1	0	0	0	0	0
	LAST	0	0	1	0	1	4	4	0	5	0	0	0
	TOTAL	0	7	3	7	2	42	5	1	12	1	0	0
YES	.200000	0	2	0	0	0	5	0	0	0	1	0	0
	.400000	0	0	0	0	0	0	1	0	5	0	0	0
	.600000	1	3	1	3	2	0	4	0	1	3	1	[illegible]
	.800000	0	0	1	0	0	0	1	0	1	0	0	0
	LAST	11	14	8	21	27	30	15	10	26	3	3	[illegible]
	TOTAL	12	19	10	24	29	35	21	10	33	7	4	7

OUT	LDAF	TYNEW 12 0000	13 0000	14 0000	15.0000	TOTAL
NO	.200000	0	0	0	0	38
	.400000	0	2	0	0	12
	.600000	0	1	0	1	18
	.800000	0	0	1	0	2
	LAST	1	0	1	0	17
	TOTAL	1	3	2	1	87
YES	.200000	1	0	0	0	9
	.400000	0	0	0	0	6
	.600000	0	3	1	0	23
	.800000	0	0	0	0	3
	LAST	6	21	8	1	211
	TOTAL	7	24	9	1	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

----- ALL-CASES HAD COMPLETE DATA FOR THIS TABLE.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	159	1137.36	.00000	2376.28	.00000
1	139	426.56	.00000	778.80	.00000
2	21	48.35	[illegible]	39.27	.00910
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	20	710.80	.00000	1597.48	.00000
2	118	378.21	.00000	739.53	.00000
3	21	48.35	.00042	39.27	.00910

NUMBER OF INTEGER WORDS OF STORAGE USED IN PRECEDING PROBLEM 3588
CPU TIME USED 20.935 SECONDS

TABLE A-6.3.2

LRFE

****
* TABLE PARAGRAPH 1 *
****

***** OBSERVED-FREQUENCY TABLE 1

OUT	LOGP	TY 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	42	5	5	6	58
	.400000	6	1	1	2	10
	.600000	2	2	3	0	7
	.800000	3	1	0	5	9
	LAST	0	1	0	2	3
	TOTAL	53	10	9	15	87
YES	.200000	28	11	5	7	51
	.400000	3	0	0	5	8
	.600000	2	1	2	0	5
	.800000	4	2	1	4	11
	LAST	48	30	34	65	177
	TOTAL	85	44	42	81	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

ALL CASES HAD COMPLETE DATA FOR THIS TABLE.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	39	674.60	.00000	1028.91	.00000
1	31	215.90	.00000	223.34	.00000
2	11	12.93	.29794	10.74	.46558
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	458.71	.00000	805.56	.00000
2	20	202.97	.00000	212.61	.00000
3	11	12.93	.29794	10.74	.46558

PAGE 13 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 2

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	8	46	1	3	58
	.400000	6	3	0	1	10
	.600000	4	1	1	1	7
	.800000	1	8	0	0	9
	LAST	0	2	0	1	3
	TOTAL	19	60	2	6	87
YES	.200000	20	16	6	9	51
	.400000	1	7	0	0	8
	.600000	2	3	0	0	5
	.800000	1	8	1	1	11
	LAST	58	65	18	36	177
	TOTAL	82	99	25	46	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

ALL CASES HAD COMPLETE DATA FOR THIS TABLE.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	39	733.59	.00000	1128.49	.00000
1	31	209.94	.00000	219.17	.00000
2	11	30.58	.00128	29.95	.00161
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	523.65	.00000	909.32	.00000
2	20	179.36	.00000	189.22	.00000
3	11	30.58	.00128	29.95	.00161

OUT	LOCP	JAN 1 00000	2 00000	4 00000	5 00000	6 00000	10 0000	11 0000	12 0000	14 0000	15 0000	17 0000	18 0[illegible]
NO	200000	0	0	12	0	0	[illegible]	32	0	0	0	0	
	400000	[illegible]	0	2	0	0	[illegible]	3	0	0	0	0	
	600000	[illegible]	0	1	0	0	[illegible]	3	0	0	0	0	
	800000	[illegible]	2	0	0	1	[illegible]	3	0	0	0	0	
	LAST	1	0	0	0	0	0	0	0	0	0	0	
	TOTAL	3	2	17	0	1	0	41	0	0	0	0	
YES	200000	0	3	5	0	0	0	1	2	1	0	2	
	400000	[illegible]	1	0	0	0	0	6	0	0	0	0	
	600000	1	0	0	0	0	0	2	0	0	0	0	
	800000	2	1	1	0	1	0	2	0	0	0	0	
	LAST	19	6	1	2	8	4	13	1	1	6	2	
	TOTAL	22	11	11	2	9	4	24	3	2	6	4	

OUT	LOCP	19 0000	21 0000	22 0000	24 0000	26 0000	BRN 27 0000	28 0000	30 0000	31 0000	32 0000	33 0000	34 0[illegible]
NO	200000	5	0	0	0	0	0	0	0	0	0	1	
	400000	[illegible]	0	0	0	0	[illegible]	0	0	0	0	0	
	600000	[illegible]	0	[illegible]	0	0	[illegible]	0	1	0	0	0	
	800000	[illegible]	0	[illegible]	0	0	[illegible]	0	0	0	0	0	
	LAST	[illegible]	[illegible]	[illegible]	0	0	[illegible]	0	0	0	0	0	
	TOTAL	[illegible]	1	[illegible]	0	0	[illegible]	0	1	0	0	[illegible]	
YES	200000	[illegible]	4	0	0	1	[illegible]	0	1	1	1	0	
	400000	[illegible]	0	0	0	0	[illegible]	0	0	0	0	0	
	600000	[illegible]	1	[illegible]	0	0	[illegible]	0	0	[illegible]	0	0	
	800000	0	0	[illegible]	0	0	[illegible]	0	1	[illegible]	0	[illegible]	
	LAST	[illegible]	[illegible]	[illegible]	11	[illegible]	[illegible]	[illegible]	3	[illegible]	[illegible]	[illegible]	
	TOTAL	4	14	[illegible]	11	4	16	1	5	11	2	3	

OUT	LOCP	35 0000	36 0000	37 0000	38 0000	39 0000	BRN 40 0000	41 0000	42 0000	43 0000	44 0000	0 45 0000	46 0
NO	200000	0	1	1	1	0	1	0	1	0	0	0	
	400000	0	0	0	0	0	0	0	0	0	0	0	
	600000	0	0	0	0	0	0	0	0	0	0	0	
	800000	0	0	0	0	0	0	0	0	0	0	0	
	LAST	0	0	0	0	0	0	0	0	0	0	0	
	TOTAL	0	1	1	1	0	1	0	1	0	0	0	
YES	200000	0	1	2	0	0	3	0	1	0	2	0	
	400000	0	0	0	0	0	0	0	0	0	0	0	
	600000	0	0	0	0	0	0	0	0	0	0	0	
	800000	0	0	0	0	0	0	0	0	0	0	0	
	LAST	4	6	3	0	1	13	5	3	6	1	2	
	TOTAL	4	7	5	0	1	16	5	4	6	3	2	

OUT	LOCP	BRN 47 0000	48 0000	TOTAL
NO	200000	1	0	58
	400000	0	0	40
	600000	1	0	[illegible]
	800000	0	0	[illegible]
	LAST	0	0	3
	TOTAL	2	0	87
YES	200000	0	1	41
	400000	0	0	8
	600000	1	0	5
	800000	0	0	11
	LAST	1	2	177
	TOTAL	2	3	242

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 329

10 CASES HAD INCOMPLETE DATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	379	1233.21	00000	3155.68	.00000
1	337	506.56	.00000	641.28	.00000
2	29	51.68	00591	59.71	.00067
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	42	726.66	.00000	2514.40	.00000
2	308	454.88	.00000	581.57	.00000
3	29	51.68	.00591	59.71	.00067

PAGE 21 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 5

OUT	LOOP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000
NO	.200000	16	7	10	7	4	2	0	2	0	0	1
	.400000	1	3	1	3	0	0	0	0	0	1	1
	.600000	1	1	3	0	1	0	0	0	0	0	0
	.800000	0	4	1	0	1	0	0	0	1	1	0
	LAST	0	0	1	1	0	1	0	0	0	0	0
	TOTAL	18	15	16	11	6	3	0	2	1	2	2
YES	.200000	0	0	2	0	0	1	0	0	0	0	0
	.400000	0	2	3	0	0	2	0	1	0	0	0
	.600000	0	0	2	3	0	0	0	0	0	0	0
	.800000	0	5	4	0	1	1	0	0	0	0	0
	LAST	6	51	57	26	17	2	4	4	1	2	0
	TOTAL	6	58	68	29	18	6	4	5	1	2	0

OUT	LOOP	AMTTN TOTAL
NO	.200000	50
	.400000	10
	.600000	7
	.800000	9
	LAST	3
	TOTAL	79
YES	.200000	3
	.400000	8
	.600000	5
	.800000	11
	LAST	177
	TOTAL	204

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 283

56 CASES HAD INCOMPLETE DATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	119	1006.79	.00000	2922.23	.00000
1	103	337.48	.00000	469.05	.00000
2	24	32.96	.10501	135.38	.00000
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	16	669.31	.00000	2453.18	.00000
2	79	304.52	.00000	333.68	.00000
3	24	32.96	.10501	135.38	.00000

PAGE [illegible] BHDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 6

OUT	LOGP	AMT 250.000	500.000	750.000	LAST	TOTAL
NO	.200000	45	7	2	4	58
	.400000	6	4	0	0	10
	.600000	5	2	0	0	7
	.800000	3	3	3	0	9
	LAST	1	1	0	1	3
	TOTAL	60	17	5	5	87
YES	.200000	42	7	1	1	51
	.400000	6	1	0	1	8
	.600000	3	1	0	1	5
	.800000	8	2	0	1	11
	LAST	135	23	12	7	177
	TOTAL	194	34	13	11	252

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 339

ALL CASES HAD COMPLETE DATA FOR THIS TABLE

PAGE 25 BMDP4F TABLES ANAL OF BDATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	39	966.85	.00000	2384.89	.00000
1	31	169.49	.00000	173.37	.00000
2	10	15.21	.12464	15.15	.12685
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR-CHISQ	PROB.	PEARSON-CHISQ	PROB.
1	8	797.36	.00000	2211.52	.00000
2	21	154.28	.00000	158.23	.00000
3	10	15.21	.12464	15.15	.12685

NUMBER OF INTEGER WORDS OF STORAGE USED IN PRECEDING PROBLEM 3318
CPU TIME USED 16.713 SECONDS

TABLE A-6.3.3

LRME

PAGE 11 BMDP TABLES ANAL OF DATA

```
**************************
* TABLE PARAGRAPH      1 *
**************************
```

***** OBSERVED FREQUENCY TABLE 1

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	32	0	0	0	32
	.400000	10	1	1	3	15
	.600000	4	2	2	4	12
	.800000	4	4	3	2	13
	LAST	0	2	0	5	7
	TOTAL	50	9	6	14	79
YES	.200000	0	1	0	0	1
	.400000	5	0	0	1	6
	.600000	2	2	0	5	9
	.800000	16	11	4	5	36
	LAST	36	20	13	43	112
	TOTAL	59	34	17	54	164

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 280

56 CASES HAD INCOMPLETE DATA

PAGE 13 BMDP4F TABLES ANAL OF CTATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO

K-FACTOR	D F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	39	582.21	.00000	915.73	.00000
1	31	282.64	.00000	385.67	.00000
2	10	22.32	.01350	51.92	.00000
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	8	299.57	.00000	530.04	.00000
2	21	260.32	.00000	333.77	.00000
3	10	22.32	.01350	51.92	.00000

PAGE 14 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 2

OUT	LPPP	NEW 1.00000	2.00000	[illegible]	4 [illegible]	TOTAL
NO	.200000	0	32	0	0	32
	.400000	6	8	0	1	15
	.600000	2	7	1	2	12
	.800000	5	8	0	0	13
	LAST	1	3	1	2	7
	TOTAL	14	58	2	5	79
YES	.200000	0	1	0	0	1
	.400000	2	3	0	1	6
	.600000	1	7	1	0	8
	.800000	7	26	2	1	36
	LAST	53	47	16	36	152
	TOTAL	63	84	19	38	204

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 283

36 CASES HAD INCOMPLETE DATA

PAGE 16 CMDF4F TABLES ANAL OF EDATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
0-MEAN	39	626.22	.000[illegible]	940.04	.00000
1	31	256.57	.000[illegible]	315.59	.00000
2	8	7.46	.488[illegible]	6.05	.64149
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	369.66	.00000	624.45	.00000
2	23	249.11	.000[illegible]	309.53	.00000
3	8	7.46	.488[illegible]	6.05	.64149

OUT	LRPP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	[illegible]	[illegible]
NO	.200000	0	[illegible]	7	0	0	0	22	0	0	0	0	[illegible]
	.400000	0	[illegible]	4	0	0	0	7	0	0	0	0	[illegible]
	.600000	[illegible]	[illegible]	0	0	0	0	[illegible]	0	0	0	0	[illegible]
	.800000	2	[illegible]	0	0	0	0	4	0	0	0	0	[illegible]
	LAST	1	[illegible]	0	0	1	0	0	0	0	0	0	[illegible]
	TOTAL	[illegible]	[illegible]	[illegible]	0	1	0	37	0	0	0	0	0
YES	.200000	0	0	1	0	0	0	0	0	0	0	0	[illegible]
	.400000	0	0	1	0	0	0	3	0	0	0	0	[illegible]
	.600000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	[illegible]
	.800000	5	2	1	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	[illegible]
	LAST	17	[illegible]	0	1	7	[illegible]	[illegible]	1	1	5	2	[illegible]
	TOTAL	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

OUT	LFPP	BRN [illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NO	.200000	3	0	0	0	0	0	0	0	0	0	0	[illegible]
	.400000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	[illegible]	[illegible]	[illegible]	[illegible]
	.600000	[illegible]	0	0	0	0	[illegible]	0	0	0	0	[illegible]	[illegible]
	.800000	[illegible]	0	0	0	0	1	0	1	0	0	0	[illegible]
	LAST	[illegible]	[illegible]	0	0	0	1	[illegible]	0	[illegible]	[illegible]	1	[illegible]
	TOTAL	[illegible]	1	0	0	0	[illegible]	0	1	0	0	[illegible]	[illegible]
YES	.200000	0	0	0	[illegible]	0	[illegible]	0	0	0	0	[illegible]	[illegible]
	.400000	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	.600000	0	0	0	0	0	0	0	0	0	0	0	[illegible]
	.800000	0	1	1	0	0	3	1	0	0	0	1	[illegible]
	LAST	1	9	0	11	3	11	0	4	10	1	2	0
	TOTAL	3	10	1	11	3	14	1	4	10	1	3	[illegible]

OUT	LFPP	BRN 35.0000	[illegible]	37.0000	[illegible]	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	[illegible]
NO	.200000	0	0	0	0	0	0	0	0	0	0	0	[illegible]
	.400000	0	0	0	[illegible]	0	0	0	0	0	0	0	[illegible]
	.600000	0	0	0	1	0	0	0	0	0	0	0	[illegible]
	.800000	0	0	1	0	0	0	0	1	0	0	0	[illegible]
	LAST	0	0	0	0	0	1	0	0	0	0	0	[illegible]
	TOTAL	0	0	1	1	0	1	0	1	0	0	0	[illegible]
YES	.200000	0	0	0	0	0	0	0	0	0	0	0	[illegible]
	.400000	0	0	0	0	0	0	0	0	0	0	0	[illegible]
	.600000	0	0	0	0	0	0	0	0	0	0	0	[illegible]
	.800000	0	0	1	0	1	1	0	1	0	0	0	[illegible]
	LAST	4	6	2	0	0	12	5	2	6	1	2	[illegible]
	TOTAL	4	6	3	[illegible]	1	13	5	3	6	1	2	[illegible]

OUT	LFPP	BRN 47.0000	49.0000	TOTAL
NO	.200000	0	0	32
	.400000	0	0	15
	.600000	2	0	12
	.800000	0	0	13
	LAST	0	0	7
	TOTAL	2	0	79
YES	.200000	0	0	[illegible]
	.400000	0	0	[illegible]
	.600000	1	0	6
	.800000	1	0	[illegible]
	LAST	0	2	152
	TOTAL	2	2	204

PAGE 17 SHEPAF TABLES ANAL ... DATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	379	1102.03	.00000	2559.61	.00000
1	337	531.20	.00000	828.11	.00000
** THE LOGLINEAR ALGORITHM DID NOT CONVERGE WHEN FITTING THE MODEL BELOW					
2	18	21.21	.2690	16.99	.52399
3	0	0.	1.	0	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANEOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	42	570.82	.0000	1731.50	.00000
2	319	510.00	.0000	811.13	.00000
3	18	21.21	.2690	16.99	.52399

PAGE 23 CNDF4T TABLES ANAL OF DDATA

***** OBSERVED FREQUENCY TABLE 5

OUT	LRPP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	13	2	8	3	3	1	0	1	0	0	1	0
	.400000	3	4	2	3	0	0	0	1	0	0	1	1
	.600000	1	[illegible]	1	2	0	0	0	0	0	1	0	1
	.800000	0	3	4	1	3	0	0	0	1	0	0	1
	LAST	1	0	1	2	0	2	0	0	0	1	0	0
	TOTAL	18	15	16	11	6	3	0	2	1	2	2	3
YES	.200000	0	0	1	0	0	0	0	0	0	0	0	0
	.400000	0	2	2	0	0	2	0	0	0	0	0	0
	.600000	0	3	[illegible]	1	[illegible]	1	0	0	0	0	0	0
	.800000	1	9	12	6	1	3	0	1	0	0	0	1
	LAST	5	44	49	19	1	4	[illegible]	4	1	2	0	5
	TOTAL	6	[illegible]	[illegible]	[illegible]	[illegible]	6	[illegible]	5	[illegible]	[illegible]	0	[illegible]

OUT	LRPP	AMTTN TOTAL
NO	.200000	32
	.400000	15
	.600000	12
	.800000	13
	LAST	7
	TOTAL	79
YES	.200000	1
	.400000	6
	.600000	9
	.800000	36
	LAST	152
	TOTAL	204

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 283

5 CASES HAD INCOMPLETE DATA

PAGE 25 BMDP4F TABLES ANAL OF BDATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	119	864.14	.00000	2189.51	.00000
1	103	288.84	.00000	441.26	.00000
2	26	40.47	.03504	83.92	.00000
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	16	575.29	.00000	1748.25	.00000
2	77	248.37	.00000	357.34	.00000
3	26	40.47	.03504	83.92	.00000

PAGE 26 BMDP4F TABLES ANAL OF BDATA

***** OBSERVED FREQUENCY TABLE 6

OUT	LPPP	AMT 250.000	500.000	750.000	LAST	TOTAL
NO	.200000	25	3	2	2	32
	.400000	10	4	[illegible]	1	17
	.600000	9	3	0	0	12
	.800000	8	3	2	0	13
	LAST	2	3	1	1	7
	TOTAL	54	16	9	4	79
YES	.200000	1	0	0	0	1
	.400000	3	3	0	0	6
	.600000	8	1	0	0	9
	.800000	20	9	5	2	36
	LAST	121	16	7	8	152
	TOTAL	153	29	12	10	204

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 283

55 CASES HAD INCOMPLETE DATA

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
0-MEAN	39	755.47	.00000	2055.52	.00000
1	31	192.98	.00000	183.89	.00000
2	9	10.36	.32196	10.20	.33484
3	0	0.	1.	0.	1.

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB	PEARSON CHISQ	PROB.
1	8	562.49	.00000	1871.63	.00000
2	22	182.61	.00000	173.69	.00000
3	9	10.36	.32196	10.20	.33484

NUMBER OF INTEGER WORDS OF STORAGE USED IN PRECEDING PROBLEM 3318
CPU TIME USED 17.545 SECONDS

TABLE A-6.4

TABLE A-6.4.1

LDA

****
*   MODEL 1   *
****

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
TL, LD.	11	26.95	0.0047	23.23	0.0164

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	36.4	0.0	0.8	0.8	38.0
	.400000	2.7	1.3	0.0	8.0	12.0
	.600000	7.5	3.5	4.8	2.2	18.0
	.800000	0.0	1.6	0.0	0.4	2.0
	LAST	5.4	3.0	2.9	5.7	17.0
	TOTAL	51.9	9.4	8.5	17.1	87.0
YES	.200000	8.6	0.0	0.2	0.2	9.0
	.400000	1.3	0.7	0.0	4.0	6.0
	.600000	9.5	4.5	6.2	2.8	23.0
	.800000	0.0	2.4	0.0	0.6	3.0
	LAST	66.6	37.0	36.1	71.3	211.0
	TOTAL	86.1	44.6	42.5	78.9	252.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.1	0.0	0.2	-0.9
	.400000	0.8	-0.3	0.0	-0.4
	.600000	0.2	-1.3	1.4	-0.8
	.800000	0.0	0.3	0.0	-0.6
	LAST	-0.6	1.7	-1.7	0.5
YES	.200000	-0.2	0.0	-0.4	1.8
	.400000	-1.2	0.4	0.0	0.5
	.600000	-0.2	1.2	-1.3	0.7
	.800000	0.0	-0.3	0.0	0.5
	LAST	0.2	-0.5	0.5	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LDAP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.1	0.0	0.4	-1.1
	.400000	0.8	-0.1	0.0	-0.3
	.600000	0.3	-1.5	1.3	-0.7
	.800000	0.0	0.4	0.0	-0.6
	LAST	-0.5	1.5	-2.6	0.6
YES	.200000	-0.1	0.0	-0.3	1.1
	.400000	-1.5	0.5	0.0	0.6
	.600000	-0.1	1.1	-1.3	0.7
	.800000	0.0	-0.1	0.0	0.6
	LAST	0.2	-0.5	0.5	-0.1

****
*   MODEL 1   *
****

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
NL, LO.	12	39.28	0.0001	40.03	0.0001

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	1.6	34.8	1.6	0.0	38.0
	.400000	2.0	8.7	0.0	1.3	12.0
	.600000	10.1	2.6	2.2	3.1	18.0
	.800000	0.4	1.2	0.0	0.4	2.0
	LAST	5.4	7.0	1.5	3.1	17.0
	TOTAL	19.5	54.3	5.3	7.9	87.0
YES	.200000	0.4	8.2	0.4	0.0	9.0
	.400000	1.0	4.3	0.0	0.7	6.0
	.600000	12.9	3.4	2.8	3.9	23.0
	.800000	0.6	1.8	0.0	0.6	3.0
	LAST	66.6	87.0	18.5	38.9	211.0
	TOTAL	81.5	104.7	21.7	44.1	252.0

PAGE 17 BMDP4F TABLES ANAL OF BDATA

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.3	0.5	-1.3	0.0
	.400000	0.7	-0.6	0.0	0.6
	.600000	1.2	-1.0	-0.8	-0.6
	.800000	-0.6	-0.2	0.0	0.9
	LAST	-1.5	2.3	-0.4	-1.2
YES	.200000	2.6	-1.1	2.6	0.0
	.400000	-1.0	0.8	0.0	-0.8
	.600000	-1.1	0.9	0.7	0.5
	.800000	0.5	0.1	0.0	-0.8
	LAST	0.4	-0.6	0.1	0.3

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LDAP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.7	0.6	-1.7	0.0
	.400000	0.7	-0.5	0.0	0.6
	.600000	1.2	-1.0	-0.7	-0.5
	.800000	-0.6	0.0	0.0	0.8
	LAST	-1.6	2.0	-0.2	-1.3
YES	.200000	1.6	-1.1	1.6	0.0
	.400000	-1.2	0.8	0.0	-0.9
	.600000	-1.1	0.9	0.7	0.6
	.800000	0.6	0.3	0.0	-0.8
	LAST	0.4	-0.6	0.2	0.4

MODEL			DF	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB	ITER.
BL,LD			46	50.30	0.3070	55.80	0.1526	2

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LDAP		BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000		0.0	0.0	6.5	0.0	0.0	0.0	25.1	0.0	0.0	0.0	0.0	0.0
	.400000		0.0	0.0	0.7	0.0	0.0	0.0	7.3	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	8.3	0.0	0.0	0.0	8.3	0.0	0.0	0.0	0.0	0.0
	.800000		0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0
	LAST		1.9	1.0	0.0	0.2	0.8	0.3	0.2	0.2	0.2	0.5	0.3	0.5
	TOTAL	-	1.9	1.0	15.5	0.2	0.8	0.3	41.4	0.2	0.2	0.5	0.3	0.5
YES	.200000		0.0	0.0	1.5	0.0	0.0	0.0	5.9	0.0	0.0	0.0	0.0	0.0
	.400000		0.0	0.0	0.3	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	10.7	0.0	0.0	0.0	10.7	0.0	0.0	0.0	0.0	0.0
	.800000		0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0
	LAST		23.1	12.0	0.0	1.8	9.2	3.7	2.8	2.8	1.8	5.5	3.7	6.5
	TOTAL	-	23.1	12.0	12.5	1.8	9.2	3.7	23.6	2.8	1.8	5.5	3.7	6.5

OUT	LDAP		BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000		6.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000		3.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4
	.800000		0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0
	LAST		0.0	1.2	0.1	0.9	0.3	1.3	0.1	0.5	0.9	0.2	0.3	0.1
	TOTAL	-	9.8	1.2	0.1	0.9	0.3	1.7	0.1	0.5	0.9	0.2	0.3	0.5
YES	.200000		1.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000		1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000		0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0
	LAST		0.0	13.8	0.9	10.1	3.7	15.7	0.9	5.5	10.1	1.8	3.7	0.9
	TOTAL	-	3.2	13.8	0.9	10.1	3.7	16.3	0.9	5.5	10.1	1.8	3.7	1.5

OUT	LDAP		BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	.200000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4
	.800000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0
	LAST		0.3	0.6	0.5	0.0	0.1	1.3	0.4	0.3	0.5	0.2	0.2	0.5
	TOTAL	-	0.3	0.6	0.5	0.4	0.1	1.3	0.4	0.7	0.5	0.2	0.2	0.9
YES	.200000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000		0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0
	LAST		3.7	7.4	5.5	0.0	0.9	15.7	4.6	3.7	5.5	2.8	1.8	5.5
	TOTAL	-	3.7	7.4	5.5	0.6	0.9	15.7	4.6	4.3	5.5	2.8	1.8	6.1

OUT	LDAP	BRN 47.0000	48.0000	TOTAL
NO	.200000	0.0	0.0	38.0
	.400000	0.7	0.0	12.0
	.600000	0.0	0.0	18.0
	.800000	0.8	0.0	2.0
	LAST	0.1	0.2	17.0
	TOTAL	1.5	0.2	87.0
YES	.200000	0.0	0.0	9.0
	.400000	0.3	0.0	6.0
	.600000	0.0	0.0	23.0
	.800000	1.6	0.0	3.0
	LAST	0.9	2.8	201.0
	TOTAL	[illegible]	2.8	242.0

STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	0.0	0.2	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.4	0.0	0.0	0.0	-0.5	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.2	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0	0.0
	LAST	0.8	1.0	0.0	-0.4	0.2	-0.6	-0.5	-0.5	-0.4	-0.7	-0.6	-0.7
YES	.200000	0.0	0.0	-0.4	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	-0.6	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	-0.2	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0
	LAST	-0.2	-0.3	0.0	0.1	-0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.2

OUT	LDAP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	.200000	-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	[illegible]
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	[illegible]
	.800000	0.0	0.0	0.0	0.0	0.0	[illegible]	0.0	0.0	0.0	0.0	0.0	[illegible]
	LAST	0.0	-0.2	-0.3	-0.9	-0.6	-0.3	-0.3	0.8	-0.9	-0.4	1.2	[illegible]
YES	.200000	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	-1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	0.0	0.0	0.0	0.0	-0.8	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.0	0.0	0.1	0.3	0.2	0.1	0.1	-0.2	0.3	0.1	-0.4	-1.0

OUT	LDAP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.7
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0
	LAST	-0.6	0.5	0.8	0.0	-0.3	-0.3	-0.6	1.2	-0.7	-0.5	-0.4	0.8
YES	.200000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.0	0.0	0.0	-0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
	LAST	0.2	-0.1	-0.2	0.0	0.1	0.1	0.2	-0.4	0.2	0.1	0.1	-0.2

OUT	LDAP	BRN 47.0000	48.0000
NO	.200000	0.0	0.0
	.400000	-0.8	0.0
	.600000	0.0	0.0
	.800000	0.2	0.0
	LAST	3.3	-0.5
YES	.200000	0.0	0.0
	.400000	1.2	0.0
	.600000	0.0	0.0
	.800000	-0.2	0.0
	LAST	-1.0	0.1

****
* MODEL 1 *
****

MODEL

LD.TL

	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
	38	66.11	0.0032	58.52	0.0178

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LDAP	TYNEW 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	11.0000
NO	.200000	0.0	1.6	0.0	0.0	0.0	34.0	0.0	0.8	0.0	0.8	0.0	0.0
	.400000	0.0	1.3	0.7	0.0	0.0	0.0	0.7	0.0	8.0	0.0	0.0	0.0
	.600000	0.4	3.5	0.9	4.4	1.3	0.4	1.8	0.0	0.4	1.8	0.4	0.0
	.800000	0.0	0.0	0.4	0.0	0.0	0.0	0.8	0.0	0.4	0.0	0.0	0.0
	LAST	0.8	1.0	0.7	1.6	2.1	2.5	1.4	0.7	2.3	0.2	0.2	6.5
	TOTAL	1.3	7.5	2.6	6.0	3.4	36.9	4.6	1.6	11.2	2.8	0.7	6.5
YES	.200000	0.0	0.4	0.0	0.0	0.0	8.0	0.0	0.2	0.0	0.2	0.0	0.0
	.400000	0.0	0.7	0.3	0.0	0.0	0.0	0.3	0.0	4.0	0.0	0.0	0.0
	.600000	0.6	4.5	1.1	5.6	1.7	0.6	2.2	0.0	0.6	2.2	0.6	0.0
	.800000	0.0	0.0	0.6	0.0	0.0	0.0	1.2	0.0	0.6	0.0	0.0	0.0
	LAST	10.2	13.0	8.3	19.4	25.9	31.5	17.6	9.3	28.7	2.8	2.8	6.5
	TOTAL	10.7	18.5	10.4	25.0	27.6	40.1	21.4	9.4	33.8	5.2	3.3	6.5

OUT	LDAP	TYNEW 12.0000	13.0000	14.0000	15.0000	TOTAL
NO	.200000	0.8	0.0	0.0	0.0	38.0
	.400000	0.0	1.3	0.0	0.0	12.0
	.600000	0.0	1.8	0.4	0.4	18.0
	.800000	0.0	0.0	0.4	0.0	2.0
	LAST	0.5	1.6	0.7	0.1	17.0
	TOTAL	1.3	4.7	1.5	0.5	87.0
YES	.200000	0.2	0.0	0.0	0.0	9.0
	.400000	0.0	0.7	0.0	0.0	6.0
	.600000	0.0	2.2	0.6	0.6	23.0
	.800000	0.0	0.0	0.6	0.0	3.0
	LAST	6.5	19.4	8.3	0.9	211.0
	TOTAL	6.7	22.3	9.5	1.5	252.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LDAP	TYNEW 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	11.0000
NO	.200000	0.0	-1.3	0.0	0.0	0.0	0.5	0.0	0.2	0.0	-0.9	0.0	0.0
	.400000	0.0	0.6	0.4	0.0	0.0	0.0	-0.8	0.0	-0.4	0.0	0.0	0.0
	.600000	-0.7	0.8	0.1	1.2	-0.3	0.8	-1.3	0.0	-0.7	-0.6	-0.7	0.0
	.800000	0.0	0.0	-0.6	0.0	0.0	0.0	0.2	0.0	-0.6	0.0	0.0	0.0
	LAST	-0.9	-1.0	0.4	-1.3	-0.8	0.9	2.2	-0.9	1.8	-0.5	-0.5	-0.7
YES	.200000	0.0	2.6	0.0	0.0	0.0	-1.1	0.0	-0.4	0.0	1.8	0.0	0.0
	.400000	0.0	-0.8	-0.6	0.0	0.0	0.0	1.2	0.0	0.5	0.0	0.0	0.0
	.600000	0.6	-0.7	-0.1	-1.1	0.2	-0.7	1.2	0.0	0.6	0.5	0.6	0.0
	.800000	0.0	0.0	0.5	0.0	0.0	0.0	-0.2	0.0	0.5	0.0	0.0	0.0
	LAST	0.3	0.3	-0.1	0.4	0.2	-0.3	-0.6	0.2	-0.5	0.1	0.1	0.2

OUT	LDAP	TYNEW 12.0000	13.0000	14.0000	15.0000
NO	.200000	-0.9	0.0	0.0	0.0
	.400000	0.0	0.6	0.0	0.0
	.600000	0.0	-0.6	-0.7	0.8
	.800000	0.0	0.0	0.9	0.0
	LAST	0.7	-1.3	0.4	-0.3
YES	.200000	1.8	0.0	0.0	0.0
	.400000	0.0	-0.8	0.0	0.0
	.600000	0.0	0.5	0.6	-0.7
	.800000	0.0	0.0	-0.8	0.0
	LAST	-0.2	0.4	-0.1	0.1

TABLE A-6.4.2

LRFE

```
*****************
*   MODEL  1    *
*****************

                                            LIKELIHOOD-RATIO            PEARSON
MODEL                                  D.F.  CHI-SQUARE    PROB    CHI-SQUARE   PROB     ITER.

TL; LO.                                 14      15.33     0.3559      12.98    0.5282      2

*****  EXPECTED VALUES USING ABOVE MODEL

OUT       LOOP                     TY
                     1.00000  2.00000  3.00000  4.00000   TOTAL

NO        .200000      37.2      8.5      5.3      6.9  !   58.0
          .400000       5.0      0.6      0.6      3.9  !   10.0
          .600000       2.3      1.8      2.9      0.0  !    7.0
          .800000       3.2      1.3      0.5      4.0  !    9.0
          LAST          0.8      0.5      0.6      1.1  !    3.0
          ---------------------------------------------------------
          TOTAL        48.5     12.7      9.8     16.0  !   87.0

YES       .200000      32.8      7.5      4.7      6.1  !   51.0
          .400000       4.0      0.4      0.4      3.1  !    8.0
          .600000       1.7      1.2      2.1      0.0  !    5.0
          .800000       3.8      1.6      0.5      4.9  !   11.0
          LAST         47.2     30.5     33.4     65.9  !  177.0
          ---------------------------------------------------------
          TOTAL        89.5     41.3     41.2     80.0  !  252.0
```

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOGP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.8	-1.2	-0.1	-0.3
	.400000	0.4	0.6	0.6	-1.0
	.600000	-0.2	0.2	0.0	0.0
	.800000	-0.1	-0.3	-0.7	0.5
	LAST	-0.9	0.7	-0.8	0.8
YES	.200000	-0.8	1.3	0.1	0.4
	.400000	-0.5	-0.7	-0.7	1.1
	.600000	0.3	-0.2	-0.1	0.0
	.800000	0.1	0.3	0.6	-0.4
	LAST	0.1	-0.1	0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.8	-1.2	-0.0	-0.3
	.400000	0.5	0.6	0.6	-0.9
	.600000	-0.1	0.3	0.2	0.0
	.800000	0.0	-0.1	-0.7	0.5
	LAST	-1.0	0.7	-0.8	0.8
YES	.200000	-0.8	1.2	0.2	0.4
	.400000	-0.4	-0.7	-0.7	1.0
	.600000	0.4	-0.0	0.1	0.0
	.800000	0.2	0.4	0.6	-0.3
	LAST	0.2	-0.0	-0.1	-0.1

```
****************
*   MODEL  1   *
****************
```

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
NL, LO.	14	43.95	0.0001	39.40	0.0003

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	14.9	33.0	3.7	6.4	58.0
	.400000	3.9	5.6	0.0	0.6	10.0
	.600000	3.5	2.3	0.6	0.6	7.0
	.800000	0.9	7.2	0.5	0.5	9.0
	LAST	1.0	1.1	0.3	0.6	3.0
	TOTAL	24.2	49.2	5.1	8.6	87.0
YES	.200000	13.1	29.0	3.3	5.6	51.0
	.400000	3.1	4.4	0.0	0.4	8.0
	.600000	2.5	1.7	0.4	0.4	5.0
	.800000	1.1	8.8	0.5	0.5	11.0
	LAST	57.0	65.9	17.7	36.4	177.0
	TOTAL	76.8	109.8	21.9	43.4	252.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-1.8	2.3	-1.4	-1.3
	.400000	1.1	-1.1	0.0	0.6
	.600000	0.3	-0.9	0.5	0.5
	.800000	0.1	0.3	-0.7	-0.7
	LAST	-1.0	0.8	-0.5	0.5
YES	.200000	1.9	-2.4	1.5	1.4
	.400000	-1.2	1.2	0.0	-0.7
	.600000	-0.3	1.0	-0.6	-0.6
	.800000	-0.1	-0.3	0.6	0.6
	LAST	0.1	-0.1	0.1	-0.1

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	NEW 1.00000	2.00000	3.00000	4.00000
NO	.200000	-2.0	2.1	-1.6	-1.4
	.400000	1.0	-1.1	0.0	0.6
	.600000	0.4	-0.8	0.6	0.6
	.800000	0.3	0.4	-0.7	-0.7
	LAST	-1.2	0.8	-0.5	0.6
YES	.200000	1.7	-2.7	1.3	1.3
	.400000	-1.3	1.1	0.0	-0.7
	.600000	-0.2	1.0	-0.6	-0.6
	.800000	0.1	-0.2	0.6	0.6
	LAST	0.2	-0.1	0.1	-0.0

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LOOP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	200000	0.0	1.8	12.3	0.0	0.0	0.0	19.3	1.2	0.6	0.0	1.2	0.6
	400000	0.0	0.6	2.2	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0
	600000	0.6	0.0	0.6	0.0	0.0	0.0	2.9	0.0	0.0	0.0	0.0	0.0
	800000	1.8	1.3	[illegible] 4	0.0	0.9	0.0	2.2	0.0	0.0	0.0	0.0	0.0
	LAST	0.3	0.1	[illegible] 0	0.0	0.1	0.1	0.2	0.0	0.0	0.1	0.0	0.1
	TOTAL -	2.7	3.8	15.6	0.0	1.0	0.1	29.7	1.2	0.6	0.1	1.2	0.7
YES	200000	0.0	1.2	8.7	0.0	0.0	0.0	13.7	0.8	0.4	0.0	0.8	0.4
	400000	0.0	0.4	1.8	0.0	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0
	600000	0.4	0.0	0.4	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0
	800000	2.2	1.6	0.6	0.0	1.1	0.0	2.8	0.0	0.0	0.0	0.0	0.0
	LAST	19.7	5.9	1.0	2.0	7.9	3.9	12.8	1.0	1.0	5.9	2.0	5.9
	TOTAL -	22.3	9.2	12.4	2.0	9.0	3.9	35.3	1.8	1.4	5.9	2.8	6.3

OUT	LOOP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	29.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	200000	4.1	2.3	0.0	0.0	0.6	1.2	0.0	0.6	0.6	0.6	0.6	0.6
	400000	2.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.6	0.6	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.4	0.0	0.0	0.0	0.0
	LAST	0.0	0.2	0.0	0.2	0.1	0.2	0.0	0.1	0.2	0.0	0.1	0.0
	TOTAL -	6.9	3.1	0.0	0.2	0.6	3.2	0.0	1.7	0.6	0.6	0.6	0.6
YES	200000	2.9	1.7	0.0	0.0	0.4	0.8	0.0	0.4	0.4	0.4	0.4	0.4
	400000	1.8	0.0	0.0	0.0	0.0	0.[illegible]	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.4	0.4	0.0	0.0	0.0	0.[illegible]	0.0	0.4	0.0	0.0	0.0	[illegible]
	800000	0.0	0.0	0.0	0.0	0.0	2.[illegible]	0.0	0.6	0.0	0.0	[illegible]	[illegible]
	LAST	[illegible]	9.8	1.0	10.8	3.0	11.8	1.0	[illegible]	9.8	[illegible]	[illegible]	[illegible]
	TOTAL -	6.1	11.9	1.0	10.8	3.4	14.8	1.0	4.0	10.2	1.4	3.4	1.4

OUT	LOOP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	200000	0.0	1.2	1.8	0.6	0.0	2.3	0.0	1.2	0.0	1.2	0.0	0.6
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.1	0.1	0.1	0.0	0.0	0.2	0.1	0.1	0.1	0.0	0.0	0.1
	TOTAL -	0.1	1.3	1.8	0.6	0.0	2.6	0.1	1.2	0.1	1.2	0.0	0.7
YES	200000	0.0	0.8	1.2	0.4	0.0	1.7	0.0	0.8	0.0	0.8	0.0	0.4
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	3.9	5.9	3.0	0.0	1.0	12.8	4.9	3.0	5.9	1.0	2.0	5.8
	TOTAL -	3.9	6.7	4.2	0.4	1.0	14.4	4.9	3.8	5.9	1.8	2.0	6.3

OUT	LOOP	BRN 47.0000	48.0000	TOTAL
NO	200000	0.6	0.6	58.0
	400000	0.0	0.0	10.0
	600000	1.2	0.0	7.0
	800000	0.0	0.0	9.0
	LAST	0.0	0.0	3.0
	TOTAL	1.8	0.6	87.0
YES	200000	0.4	0.4	41.0
	400000	0.0	0.0	8.0
	600000	0.8	0.0	5.0
	800000	0.0	0.0	11.0
	LAST	1.0	2.0	177.0
	TOTAL	2.2	2.4	242.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOCP	BRN 1.00000	2.00000	4.00000	5.00000	6.00000	10.0000	11.0000	12.0000	14.0000	15.0000	17.0000	18.0000
NO	.200000	0.0	-1.3	-0.1	0.0	0.0	0.0	2.9	-1.1	-0.8	0.0	-1.1	-0.8
	.400000	0.0	-0.7	1.2	0.0	0.0	0.0	-0.9	0.0	0.0	0.0	0.0	0.0
	.600000	-0.8	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.800000	0.1	0.6	-0.7	0.0	0.1	0.0	0.5	0.0	0.0	0.0	0.0	0.0
	LAST	1.2	-0.3	-0.1	-0.2	-0.4	-0.3	-0.5	-0.1	-0.1	-0.3	-0.2	-0.3
YES	.200000	0.0	1.6	0.1	0.0	0.0	0.0	-3.4	1.3	0.9	0.0	1.3	0.9
	.400000	0.0	0.8	-1.3	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.9	0.0	-0.6	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	0.0	0.0
	.800000	-0.1	-0.5	0.6	0.0	-0.1	0.0	-0.5	0.0	0.0	0.0	0.0	0.0
	LAST	-0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0

OUT	LOCP	BRN 19.0000	21.0000	22.0000	24.0000	26.0000	27.0000	28.0000	30.0000	31.0000	32.0000	33.0000	34.0000
NO	200000	0.4	-1.5	0.0	0.0	-0.8	-1.1	0.0	-0.8	-0.8	-0.8	0.5	0.5
	.400000	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	.600000	0.5	-0.8	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	-0.7	0.0	0.0	0.0	0.0
	LAST	-0.1	2.0	-0.1	-0.4	-0.2	1.8	-0.1	-0.2	-0.4	-0.1	-0.2	-0.1
YES	200000	-0.5	1.8	0.0	0.0	0.9	1.3	0.0	0.9	0.9	0.9	-0.6	-0.6
	.400000	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	-0.6	0.9	0.0	0.0	0.0	0.0	0.0	-0.6	0.0	0.0	0.0	0.0
	.800000	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.6	0.0	0.0	0.0	0.0
	LAST	0.0	-0.3	0.0	0.1	0.0	-0.2	0.0	0.0	0.1	0.0	0.0	0.0

OUT	LOCP	BRN 35.0000	36.0000	37.0000	38.0000	39.0000	40.0000	41.0000	42.0000	43.0000	44.0000	45.0000	46.0000
NO	200000	0.0	-0.2	-0.6	0.5	0.0	-0.9	0.0	-0.2	0.0	-1.1	0.0	-0.5
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	-0.3	-0.3	-0.2	0.0	-0.1	-0.5	-0.3	-0.2	-0.3	-0.1	-0.2	-0.3
YES	200000	0.0	0.2	0.7	-0.6	0.0	1.0	0.0	0.2	0.0	1.3	0.0	-0.6
	400000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	600000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	800000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	LAST	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0

OUT	LOCP	BRN 47.0000	48.0000
NO	.200000	0.5	-0.8
	.400000	0.0	0.0
	.600000	-0.2	0.0
	.800000	0.0	0.0
	LAST	-0.1	-0.2
YES	200000	-0.6	0.9
	.400000	0.0	0.0
	.600000	0.2	0.0
	.800000	0.0	0.0
	LAST	0.0	0.0

MODEL 1

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
LO, AL	36	46.92	0.1052	52.84	0.0348

EXPECTED VALUES USING ABOVE MODEL

	LOGP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	15.1	6.6	11.3	6.6	3.8	2.8	0.0	1.9	0.0	0.0	0.9	0.9
	.400000	0.6	2.8	2.2	1.7	0.0	1.1	0.0	0.6	0.0	0.6	0.6	0.0
	.600000	0.6	0.6	2.9	1.7	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	4.0	2.2	0.0	0.9	0.5	0.0	0.0	0.5	0.5	0.0	0.5
	LAST	0.1	0.9	1.0	0.5	0.3	0.1	0.1	0.1	0.0	0.0	0.0	0.1
	TOTAL	16.3	14.9	19.7	10.5	5.5	4.4	0.1	2.5	0.5	1.0	1.5	2.1
YES	.200000	0.9	0.4	0.7	0.4	0.2	0.2	0.0	0.1	0.0	0.0	0.1	0.1
	.400000	0.4	2.2	1.8	1.3	0.0	0.9	0.0	0.4	0.0	0.4	0.4	0.0
	.600000	0.4	0.4	2.1	1.2	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.4
	.800000	0.0	5.0	2.7	0.0	1.1	0.5	0.0	0.0	0.5	0.5	0.0	0.5
	LAST	5.9	50.1	57.0	26.5	16.7	2.9	3.9	3.9	1.0	2.0	0.0	6.9
	TOTAL	7.7	58.1	64.3	29.5	18.5	4.6	3.9	4.5	1.5	3.0	0.5	7.9

OUT	LOGP	AMTTN TOTAL
NO	.200000	50.0
	.400000	10.0
	.600000	7.0
	.800000	9.0
	LAST	3.0
	TOTAL	79.0
YES	.200000	3.0
	.400000	8.0
	.600000	5.0
	.800000	11.0
	LAST	177.0
	TOTAL	204.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LOOP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.2	0.2	-0.4	0.2	0.1	-0.5	0.0	0.1	0.0	0.0	0.1	0.1
	.400000	0.6	0.1	-0.8	1.0	0.0	-1.1	0.0	-0.7	0.0	0.6	0.6	0.0
	.600000	0.5	0.5	0.0	-1.3	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5
	.800000	0.0	-0.0	-0.8	0.0	0.1	-0.7	0.0	0.0	0.8	0.8	0.0	0.8
	LAST	-0.3	-0.9	0.0	0.8	-0.5	4.2	-0.3	-0.3	-0.1	-0.2	0.0	-0.3
YES	.200000	-1.0	-0.6	1.6	-0.6	-0.5	2.0	0.0	-0.3	0.0	0.0	-0.2	-0.2
	.400000	-0.7	-0.1	0.9	-1.2	0.0	1.2	0.0	0.8	0.0	-0.7	-0.7	0.0
	.600000	-0.6	-0.6	-0.1	1.6	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	-0.6
	.800000	0.0	0.0	0.8	0.0	-0.1	0.6	0.0	0.0	-0.7	-0.7	0.0	-0.7
	LAST	0.0	0.1	-0.0	-0.1	0.1	-0.6	0.0	0.0	0.0	0.0	0.0	0.0

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	AMTTN 0.00000	1.00000	2.00000	3.00000	4.00000	5.00000	6.00000	7.00000	8.00000	9.00000	10.0000	LAST
NO	.200000	0.3	0.2	-0.3	0.2	0.2	-0.4	0.0	0.2	0.0	0.0	0.2	0.2
	.400000	0.6	0.3	-0.7	1.0	0.0	-1.3	0.0	-0.8	0.0	0.6	0.6	0.0
	.600000	0.6	0.6	0.2	-1.8	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	.800000	0.0	0.1	-0.7	0.0	0.3	-0.7	0.0	0.0	0.7	0.7	0.0	0.7
	LAST	-0.2	-1.1	0.2	0.7	-0.5	1.3	-0.1	-0.1	-0.0	-0.1	0.0	-0.2
YES	.200000	-1.2	-0.6	1.2	-0.6	-0.4	1.1	0.0	-0.2	0.0	0.0	-0.1	-0.1
	.400000	-0.7	0.0	0.9	-1.5	0.0	1.0	0.0	0.7	0.0	-0.7	-0.7	0.0
	.600000	-0.6	-0.6	0.1	1.3	-0.6	0.0	0.0	0.0	0.0	0.0	0.0	-0.6
	.800000	0.0	0.1	0.8	0.0	0.1	0.6	0.0	0.0	-0.8	-0.8	0.0	-0.8
	LAST	0.1	0.2	0.0	-0.1	0.1	-0.4	0.1	0.1	0.2	0.2	0.0	0.1

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
LO, AL	13	19.17	0.1180	19.85	0.0990

***** EXPECTED VALUES USING ABOVE MODEL

OUT[illegible]	LOOP	AMT 250,000	500,000	750,000	LAST	TOTAL
NO	200000	46.3	7.4	1.6	2.7	58.0
	400000	6.7	2.8	0.0	0.6	10.0
	600000	4.7	1.7	0.0	0.6	7.0
	800000	5.0	2.2	1.3	0.5	9.0
	LAST	2.3	0.4	0.2	0.1	3.0
	TOTAL	64.8	14.6	3.1	4.4	87.0
YES	200000	40.7	6.6	1.4	2.3	51.0
	400000	5.3	2.2	0.0	0.4	8.0
	600000	3.3	1.2	0.0	0.4	5.0
	800000	6.0	2.7	1.6	0.6	11.0
	LAST	133.7	23.6	11.8	7.9	177.0
	TOTAL	189.2	36.4	14.9	11.6	252.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL -

OUT	LOGP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.2	-0.2	0.3	0.8
	.400000	-0.3	0.7	0.0	-0.7
	.600000	0.2	0.2	0.0	-0.8
	.800000	-0.9	0.5	1.4	-0.7
	LAST	-0.8	0.9	-0.4	2.4
YES	.200000	0.2	0.2	-0.3	-0.9
	.400000	0.3	-0.8	0.0	0.8
	.600000	-0.2	-0.2	0.0	0.9
	.800000	0.8	-0.5	-1.3	0.6
	LAST	0.1	-0.1	0.1	-0.3

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LOGP	AMT 250.000	500.000	750.000	LAST
NO	.200000	-0.2	-0.1	0.4	0.8
	.400000	-0.2	0.8	0.0	-0.8
	.600000	0.3	0.3	0.0	-0.8
	.800000	-0.8	0.6	1.2	-0.7
	LAST	-0.8	0.8	-0.3	1.2
YES	.200000	0.2	0.3	-0.2	-0.8
	.400000	0.4	-0.7	0.0	0.7
	.600000	-0.1	-0.0	0.0	0.8
	.800000	0.8	-0.3	-1.8	0.6
	LAST	0.1	-0.1	0.1	-0.2

TABLE A-6.4.3

LRME

****
*   MODEL 1   *
****

MODEL	D.F.	LIKELIHOOD-RATIO CHI-SQUARE	PROB	PEARSON CHI-SQUARE	PROB
TL, LO.	13	23.26	0.0387	43.53	0.0000

***** EXPECTED VALUES USING ABOVE MODEL

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000	TOTAL
NO	.200000	31.0	1.0	0.0	0.0	32.0
	.400000	10.7	0.7	0.7	2.9	15.0
	.600000	3.4	2.3	1.1	5.1	12.0
	.800000	5.3	4.0	1.9	1.9	13.0
	LAST	1.6	1.0	1.5	3.0	7.0
	TOTAL	52.1	8.9	5.2	12.9	79.0
YES	.200000	1.0	0.0	0.0	0.0	1.0
	.400000	4.3	0.3	0.3	1.1	6.0
	.600000	2.6	1.7	0.9	3.9	9.0
	.800000	14.7	11.0	5.1	5.1	36.0
	LAST	34.4	21.0	31.5	65.0	152.0
	TOTAL	56.9	34.1	37.8	75.1	204.0

***** STANDARDIZED DEVIATES = (OBS - EXP)/SQRT(EXP) FOR ABOVE MODEL

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.2	-1.0	0.0	0.0
	.400000	-0.2	0.3	0.3	0.1
	.600000	0.3	-0.2	0.8	-0.5
	.800000	-0.6	0.0	0.8	0.1
	LAST	-1.3	1.0	-1.2	1.2
YES	.200000	-1.0	5.6	0.0	0.0
	.400000	0.3	-0.5	-0.5	-0.1
	.600000	-0.4	0.2	-0.9	0.6
	.800000	0.3	-0.0	-0.5	-0.1
	LAST	0.3	-0.2	0.3	-0.2

***** FREEMAN-TUKEY DEVIATES FROM THE ABOVE MODEL

OUT	LRPP	TY 1.00000	2.00000	3.00000	4.00000
NO	.200000	0.2	-1.2	0.0	0.0
	.400000	-0.1	0.5	0.5	0.2
	.600000	0.4	-0.0	0.8	-0.4
	.800000	-0.5	0.1	0.8	0.2
	LAST	-1.7	0.9	-1.6	1.1
YES	.200000	-1.2	1.4	0.0	0.0
	.400000	0.4	-0.5	-0.5	0.1
	.600000	-0.2	0.3	-1.1	0.6
	.800000	0.4	0.1	-0.4	0.0
	LAST	0.3	-0.2	0.3	-0.2